主编

郭起浩　洪　震

神经心理评估

（第二版）

SHENJING XINLI

P I N G G U

上海科学技术出版社

Shanghai Scientific & Technical Publishers

神经心理评估

（第二版）

主编　郭起浩　洪　震

上海科学技术出版社

图书在版编目(CIP)数据

神经心理评估/郭起浩，洪震主编. —2版. —上海：上海科学技术出版社，2016.2
ISBN 978-7-5478-2936-3

Ⅰ.①神… Ⅱ.①郭… ②洪… Ⅲ.①神经心理学—心理测验—评估 Ⅳ.①R741

中国版本图书馆CIP数据核字(2015)第317383号

神经心理评估(第二版)
主编 郭起浩 洪 震

上海世纪出版股份有限公司
上 海 科 学 技 术 出 版 社 出版
(上海钦州南路71号 邮政编码200235)
上海世纪出版股份有限公司发行中心发行
200001 上海福建中路193号 www.ewen.co
常熟市华顺印刷有限公司印刷
开本787×1092 1/16 印张26.5
字数425千字
2013年1月第1版
2016年2月第2版 2016年2月第3次印刷
ISBN 978-7-5478-2936-3/R·1054
定价：58.00元

内容提要

本书在第一版介绍目前国外常用、临床医师急需了解的30多种神经心理测验的基础上，扩充了近一倍的测验，力求反映针对老年认知障碍的临床神经心理测验的新进展，内容包括测验的来源、版本演变、操作步骤、信度、效度、正常值、机制研究进展、临床意义和使用注意事项。掌握这些测验有助于临床诊断、治疗效果评估、康复措施制定，对于了解大脑的信息处理机制也有重要意义。

本书内容详细、新颖、实用，适合神经科、精神科、老年科、心理科的临床医师、研究人员与护理人员，以及有志于从事认知障碍研究的研究生、对脑科学与行为科学有兴趣的读者。

作 者 名 单

主　编

郭起浩　洪　震

编写者（以姓氏笔画为序排列）

马　洁　王　萍　汤　乐　杨　青

张曙映　陈科良　周　燕　赵倩华

洪　震　徐　岩　郭起浩　梁小妞

再版前言

《神经心理评估》第一版于2013年1月出版，令人惊喜的是，本书在业界颇受欢迎，常常有师长、同道、朋友反映本书内容新颖实用，为从事老年脑科学诊疗实践与研究的人们提供了入门读物。但也有不少认真的读者反映，本书还不够系统全面，有些临床工作中常用的测验或量表还没有收入书中。于是，我们再接再厉，扩充内容，修订成新版。

在本版中，我们力求反映针对老年认知障碍的临床神经心理测验的新进展，如收入了我们引进与翻译的蒙特利尔认知评估基础量表(MoCA-B)与社会认知评估的3个常用测验，这些测验在中国人群中的信度效度与正常值尚在验证中。为了比较全面地反映各种认知领域的不同成分，我们在记忆、注意、空间与执行功能检测中增加了10余个经典测验，大大丰富了对特定认知领域的了解与评估。因为本书主要围绕老年认知功能障碍患者的评估，所以，在照料者与知情者评估方面增加了4个国际常用的量表。此外，对有关情绪评估与计算机辅助评估的内容也做了扩充。轻度认知损害(MCI)的研究正在走向深入，作为诊断核心的标准化神经心理测验是当前的一个研究热点，我们不仅针对阿尔茨海默病所致MCI、帕金森病所致MCI、血管性MCI的专家共识的相关内容进行深入解读，也提出基于中国文化背景的成套测验方案，希望这些思考与探索有益于国内同道的工作。在参考文献部分，增加了近3年我们团队发表的SCI论文20余篇的目录。从总数127篇的论文目录中，可以看到我们一直在努力、在进步，但广度与速度依然逊色。

有些国内外常用的量表我们没有介绍，比如国内已经有的比较好的基础

的失语症检测量表，有高素荣、王荫华等教授编撰的鸿篇巨著中的内容，就不在本书重复了；国内外许多单位采用剑桥老年认知检查（CAMDEX）与 Addenbrookes 认知检查系列版本（ACE、ACE－R、ACE－Ⅲ），这两种筛查量表就其内容来说，与第三章认知筛查量表中已经介绍的几种大同小异，只能割爱。我们也没有介绍金字塔与棕榈树测验、EXIT25、汉诺塔测验（tower of Hanoi，TOH）、谚语理解测验、钉板测验、EC301 数字与计算测验等，这些测验我们曾经使用或发表过论文，但国际上目前很少使用，文献不多，若有读者感兴趣，可以直接写邮件向我们索取。

神经心理学的发展日新月异，尤其是计算机技术、互联网技术、影像学技术的进一步应用，使神经心理领域不断涌现新理论、新观点、新方法、新面貌，如自动语言分析、在线认知激活、互联网＋测验等，这些技术齐头并进，揭示人类大脑的奥秘。我们希望与时俱进，在大约 5 年后修订的第三版中会对此方面的进展做详细介绍。我们在与北京师范大学毕彦超教授、韩在柱教授合作研究语义记忆的脑机制的过程中，也积累了大量的测验数据，获得了一系列创新性成果，这些方法与结果也将在第三版中再做全面介绍。

在本版的修订过程中，我们要感谢复旦大学附属华山医院神经内科董强主任及卢家红、王毅、王坚、赵重波副主任的鼎力支持，感谢我们认知组成员赵倩华副教授与周燕主管护师长年累月的付出与不辞辛劳的分担，感谢博士研究生王萍、陈科良、杨青、汤乐和硕士研究生徐岩、梁小妞为此倾注的大量心血，感谢同济大学医学院张曙映副教授与上海中医药大学附属岳阳中西医结合医院马洁医师的贡献。从广泛阅读文献、密切联系国内临床实践，一直到严谨细致的校对，如果没有他(她)们的辛勤劳动，就没有本书修订的顺利完成。

最后，我们要感谢这些测验的原作者与使用者，每个测验有许多不同版本、有几百篇文献，因为篇幅限制，我们没有注明每个测试结果与观点的来源，没有列举大量参考文献的附录，只能选择最主要的百余篇文献附在书后，我们要向这些文献的作者表示深深的歉意与谢意。

郭起浩　洪　震

2015 年 10 月

第一版前言

随着我国社会人口的老龄化，老年人口比例在不断增加，医院门急诊的老年就诊者增多，几乎一半的老年人会有记忆、语言等认知缺损，针对这些认知功能缺损的不同领域与不同严重度的评估手段，依靠神经心理学，特别是神经心理测验的发展，神经心理测验的作用，包括认知障碍、情绪障碍和行为障碍的评估，是获得患者心理与行为变化定量信息的一种方法，对患者康复措施的制定、治疗效果的判断，鉴别早期患者与正常衰老具有重要作用，尤其是那些还没有影像变化和临床表现的早期患者，神经心理评估敏感性较高，可以检测出一般临床检查不能发现的行为、功能障碍。欧美国家已经有大量系统的专著问世，但是国内出版的实用的中文专门书籍还是非常少。

文章千古事，得失寸心知。在本书即将付梓之际，笔者认为有必要澄清以下问题。

首先，是涉及版权或者知识产权的问题。本书介绍的认知功能评估方法大部分是从国外引进的，不少读者会问，如果我采用这些测验，是否需要支付版权费或存在侵权行为？我们在编写本书时十分注意这个问题。韦氏智力测验(WIS)第Ⅰ～Ⅳ版与韦氏记忆测验(WMS)第Ⅰ～Ⅲ版是需要购买版权的测验，我们没有在本书中做详细介绍。本书介绍的测验都是可以自由地被读者在临床与科研中采用的，这些评估方法要么是因为年代久远超过产权保护期限的，如连线测验(TMT)，迄今已经有80多年使用历史；要么是编制者已经公开在期刊或网络上发表并希望读者应用的，如蒙特利尔认知评估量表(MoCA)。但是，有志于从事神经心理测验研究的同道需要注意的是，在应用

这些工具完成的论文中应注明采用的是哪位研究者编制、修订或翻译的版本，其常模和划界分需要考虑地区差异等。国内有不少论文在方法介绍中忽略他人的劳动，不注明自己采用的测验或量表的出处，或者只注明英文原版的出处而不注明中文版本来源，这应该说已涉及学术规范。本书中有部分测验是笔者首先翻译并验证的、部分数据是笔者临床使用情况的总结，当然也希望在被转载或引用时予以注明。

其次，本书没有全面介绍神经心理评估。英文的神经心理评估专著如Strauss等编著的《神经心理测验概略》(*A compendium of Neuropsychological Tests*，2006年第3版)详细介绍了100多种常用测验的使用方法、品质、常模和作用，有1 200多页。Lezak编写的《神经心理测评》(*Neuropsychological Assessment*，2004年第4版)中简要介绍的测验约500种，有1 000多页。相比这些英文专著，本书的篇幅要小得多，介绍的测验领域也不全面(如病前智力、感知觉功能、运动功能、计算功能、适应功能、成就评估、司法神经心理评估、艺术欣赏的脑机制等众多领域没有涉及)，这是因为编写本书的基本原则是，围绕老年人的轻度认知损害(MCI)与血管性认知损害(VCI)的诊断与识别，选择国际公认的、我们已有使用经验与数据积累的测验。当然，这些测验也是经典的神经心理测验的核心部分。国内尚未引进的测验不在本书的范围内。另外，作为神经心理评估的一部分，情绪与行为评估已经有精神科医生编写的专著，所以，这方面的内容也仅仅局限在常用的几个量表的介绍。新的、探索性的神经心理评估方法层出不穷，希望数年后本书第二版时能够在广度与深度方面有新的突破。

再次，测验结果的解释不应简单化和绝对化。初学者常见的错误是将测验得分或将一个测验工具的评估结果等同于临床诊断。一个好的测验有比较好的信度(reliability)和效度(validity)：信度指测量的稳定性和一致性，包括一致性信度和再测信度；效度指测验能够测量想要测量的东西，包括内容效度、效标效度和构想效度。也有比较好的敏感性(sensitivity)和特异性(specificity)：敏感性是筛查测验发现真正的病例的能力；特异性是该测验确定真正的非病例即正常人的能力。但是，测验误差的来源不仅仅是测验本身的品质，还有可能是翻译的误差，文化背景的差异，受试者的年龄、性别、教育

程度、社会经济状况的差异，测试时的心理状态(如应激、焦虑或抑郁导致注意力不集中)，施测者的技术水平和评估时机等复杂的因素。所以，测验得分只能是诊断的辅助工具，就像影像学对于脑肿瘤的诊断价值一样。另外，我们不仅要看受试者最后的得分，还要观察分析受试者的信息加工过程、各部分得分的分布。比如，笔者于2002年11月发现国内第一例语义性痴呆(SD)患者，便是因为该患者就诊时简明精神状态量表(MMSE)评估还有18分的情况下已经不能命名钢笔、手表等简易物品，从这一反常现象的深入细致的评估开始，我们与认知心理学家合作开发了一系列的语义记忆检测方法，初步的研究结果已经发表在国际杂志上。迄今我们已经积累从极轻到极重的SD患者40多例，并获得国家自然科学基金的支持对其特征性认知缺损进行深入研究。

本书的编写，要感谢吕传真教授和张明园教授的不倦教诲和指导；感谢香港中文大学心理学系陈瑞燕教授、北京师范大学心理学系毕彦超教授和韩在柱教授、安徽医科大学汪凯教授、日本京都大学手良相教授的热情支持与帮助；感谢神经内科同事董强、吴志英、卢家红、王毅、丁玎、赵倩华、虞培敏、周燕等，他们的关心与践行促进了神经心理测验在本院的发展与壮大；感谢已经毕业及在读的神经心理学方向的研究生袁晶、金丽琳、孙一忞、陈美蓉、曹歆轶、王蓓、师云波、马洁，他们为神经心理评估的应用与开发、数据的积累与统计付出了大量劳动；感谢我们的朋友李放博士、张曙映博士、葛沁昭小姐的鼎力支持与协助。最后，我们要感谢上海科学技术出版社的编辑萧烈女士的策划与鼓励。

本书在繁忙的临床工作之余编写，加上笔者对临床神经心理评估的认识水平有限，不足之处在所难免，敬请广大读者与同仁不啬指教。

2012年6月于复旦大学附属华山医院

郭起浩　洪　震

目　录

第一章

认知的大脑定位和神经心理障碍

神经心理学是研究脑与心理相互关系的学科，其历史从 Carl Wemicke (1874)开始。在早期阶段，神经心理学是神经病学(相当于硬件)与心理学(相当于软件)的结合。现代神经心理学是认知科学与临床神经科学的整合，是脑科学的主要分支之一。行为神经病学、临床神经心理学和神经精神病学构成神经心理学的主要组成部分。

在认知的大脑定位研究早期，人们认为可以发现所有认知功能处在大脑中的特定位置，简洁明快的点对点的定位思想占了主导地位。自从 Lashley 通过动物实验提出脑损坏容积起作用的观点后，大脑每个结构功能相同的思想与严格的脑结构定位观念分庭抗礼。随着资料的积累和认识的深化，人们发现有多个脑结构联合执行某种认知功能，比较著名的有 Luria 的 3 个基本功能联合区学说和 Damasio 的 3 个系统学说。Luria 的基本理论是：第一联合区由大脑半球内侧部皮质、间脑、脑干网状结构组成，调节机体紧张度，保证觉醒状态。第二联合区由枕叶视区、颞叶听区、顶叶普通感觉区组成，作用是兴奋的接受、传递、加工和储存。第三联合区由前额叶组成，对传入信息进行第二次加工、形成自己的行动计划和程序、在言语参与的条件下调控自己的行为。Damasio 的基本理论是：初级皮质记录感觉和运动信息；辅助运动区、边缘结构(内嗅皮质、海马、杏仁核和扣带皮质)、新纹状体和小脑等结构组成内部信息编码的基础；在这两者间有正反馈和负反馈联系；丘脑、前脑底部、下丘脑、脑干组成伺服系统。

从认知功能出发寻求大脑结构定位的思路与现代功能影像学的设计原理不谋而合。每个认知亚成分可能在大脑神经网络中存在许多相关结点，如 Mesulam 提出的记忆网络模型中，将记忆的感知、编码和提取区分为“外显记

忆”、“工作记忆”、“面孔识别”和“空间意识”等不同成分后，认为外显记忆的结点在海马-内嗅区域，工作记忆-筹划执行功能的结点在前额叶外侧皮质，面孔与物品再认的结点在颞中和颞极皮质，空间意识的结点在后顶皮质。尽管记忆的编码和提取涉及几乎所有的联络皮质，但其信息处理过程存在有规律的解剖分布，相关的皮质区域对感觉信息进行编码，边缘系统将这些感觉信息整合到原有的知识体系中，前额叶指导信息储存和提取的条理性。

第一节　运用的相关脑区和失用症

一、运用的相关脑区

运用是个体生活中经过模仿、学习和长期实践而形成的，是运动功能的高级表现形式，为人类所特有。作为肌群的运动基础，有 4 个相关脑结构通路，包括皮质脊髓束、皮质脑干束、内侧纵束和红核脊髓束。行动计划又称运动程序，指意向性运动的抽象表现，包括达到目标的一系列动作是如何动、动的次序、动的方向、动的时限等，这些信息产生依赖初级运动皮质外的脑结构，包括辅助运动区、前扣带回皮质、额叶眼区和顶叶。这些大脑结构与协调的、熟练的运用功能相关。

1. 辅助运动区(supplementary motor area，SMA)　在运动的计划、准备和开始方面起重要作用。SMA 位于初级运动皮质前侧，并与之相连，经常被称为前运动区。不同于初级运动皮质，每侧 SMA 有同侧和对侧运动皮质投射。功能影像学研究表明需要复杂运动排序时双侧 SMA 被激活，而在简单的重复性任务中，尽管对侧初级运动皮质区域的血流量增加了，SMA 并没有被激活。要求患者想象而不是实际执行复杂的手指序列任务时，SMA 被激活，而初级运动皮质区域没有被激活。

2. 前扣带回皮质　位于扣带沟之下、胼胝体之上，在运动和认知功能的连接中起作用，尤其是这种连接是新的或需要更多认知控制的时候。它只有 5 个细胞层，是大脑中最原始的结构之一。在分别需要手部、口部和眼睛运动的 PET 研究中，每种运动区分为熟练动作和新学习动作两种，发现只有在新学习动作中，需要手部运动时前扣带回皮质的尾部被激活；需要眼睛运动时前扣带回皮质前端被激活；需要言语时前扣带回皮质的中段被激活。

3. 额叶视野(frontal eye fields)　额叶视野位于SMA的前侧、Broca区的背面,在控制随意的眼球运动中起作用。在突闻大声或明亮物品移动时吸引人们注意引起的反射性眼睛运动是不随意的,这种运动受控于上丘。而随意的眼睛运动是编程的,如在熙熙攘攘的人群中寻找朋友的容颜,受控于额叶视野。尽管在最后的输出通路上相同,随意和不随意眼睛运动两者相互独立,但随意眼睛运动优先输出。额叶视野在成对眼睛运动即眼睛侧向快速运动中起重要作用。

4. 顶叶　顶叶在运动程序中的作用是双重的,首先涉及各种空间运动控制,其次是产生复杂的、熟悉的运动行为。前者依赖于顶叶上部区域,后者依赖于顶叶下部区域。顶叶既对自身各部分位置的本体感受信息很敏感,也接受运动和运动前区的反馈,两者整合确保身体各部服从指挥。顶叶上区的损伤可导致患者运用肢体从事自控动作的能力下降,顶叶下区的损伤会导致复杂的、熟练掌握的运动行为受影响,引起失用。

二、失用症

1. 定义　失用症(apraxia)是指在具有健全的肌力和完整的神经支配的情况下,机体不能顺利完成有目的的动作,丧失已获得的、熟练的正常运动。这种丧失不能用肌力减退、肌张力异常、震颤或舞蹈症等解释,它是皮质高级运动损害而不是初级水平的运动控制损害。失用的表现一般是双侧的,而肢体肌肉控制障碍时病灶多为单侧。失用患者有时能自发地完成熟练运动动作,他们的困难仅出现在要求完成有目的动作或模仿无意义动作和不熟悉姿势的时候。

2. 失用症的检查　失用症的检查项目见表1-1-1。每个项目又可分为自发完成和模仿完成两个步骤。

表1-1-1　失用症的检查项目

部　位	动　　作
上肢	不及物:表示再见、敬礼、搭便车、来去、停止 及物:用钥匙开门、抛硬币、使用剪刀、敲音叉、用订书机装订、开调料瓶、使用螺丝起子、使用锤子 描述:描三角形、圆形、螺旋形 系列动作:折好信纸、放入信封、封好信封口、贴上邮票
下肢	伸腿、踢球动作、踩灭烟头动作、走直线、马步、后退

（续表）

部 位	动 作
口部	不及物动作：伸舌、露齿、张口、吹口哨、发鞭炮声（“噼啪”） 及物动作：吹灭火柴、吮吸管
眼睛	闭眼、左看、右看、上视、眯眼
全身	鞠躬、扫地动作、转圈子、给自行车打气动作

3. 分类　失用有许多分类方法。以部位分，可以区分为口面失用和肢体失用。前者定位于额叶和中央顶盖、邻近的颞上回和岛回的前端，后者与左顶叶或顶颞叶区病变有关。1905 年 Liepmann 将失用症分为观念性失用（ideational apraxia）和观念运动性失用（ideomotor apraxia），前者是形成运动的“意念或观念”能力受损，比如点蜡烛，需要取火柴、关火柴盒、擦火、火头凑上蜡烛等连贯动作，患者能完成单个动作，但不能组织每个步骤成为系列动作。另外，患者不能使用锤子、牙刷、剪刀等实际物品。后者是概念和行动之间脱节，信息不能从运动的大脑区域传输到指挥身体进行运动的区域，与前者不同，患者不能完成或模仿简单具体的动作，如手势和不及物动作如敬礼，但能完成复杂动作的次序组织。有些研究者使用“符号性动作失用”和“物品使用的失用”来取代意念性失用和意念运动性失用。其他类型的失用如结构失用（不能正确处理空间关系）和穿衣失用都是空间领域的失用，与空间加工和偏侧忽视有关，许多神经心理学家将它们归到视觉空间障碍。

第二节　物体识别与失认症

一、物体识别的理论

物体识别看起来非常简单和自然，但事实上其加工过程相当复杂。Marr 提出物体识别的计算理论，首先定义 3 种主要表征（即符号性描述）：初级简图，是对视觉输入的边缘、亮暗和轮廓的二维描述；$2\frac{1}{2}$-D 简图，通过利用阴影、纹理、运动和双眼视差等信息对可视表面深度和方位进行描述；3-D 模型表征：这一表征描述物体形状的三维特征及它们的相对位置。初级简图和 $2\frac{1}{2}$-D简图围绕视网膜的参照系统展开，因而它们保留观察者和被观察物体

之间关系的信息。与前两种表征不同，3-D模型表征的参照系主轴是由物体特性决定的，是以客体为中心的，不依赖于观察者的角度变化，具有物体认知的恒常性，所以，尽管视网膜显示的成像不同，我们仍然感知到是同一物体。该理论对于我们细致地理解各种辨认障碍提供了有用的框架。综合各种物体识别理论(模板理论、特征理论和成分识别理论)，以下一系列加工对于物体识别是重要的：对物体边缘形状的编码、物体各种特征的组合、与大脑储存的结构性知识匹配及语义知识的提取。

二、失认症

1. 定义　失认症(agnosia)是物体识别的神经心理障碍。

2. 视觉性失认症的分类　视觉性失认分为知觉性视觉失认和联合性视觉失认。前者因严重的知觉加工缺陷引起物体识别困难；后者为知觉加工完整，物体识别的困难是因对目标的视觉性记忆损害或不能搜索到与目标相关的语义信息所致。

(1) 知觉性视觉失认：可分为狭义知觉性失认、同时性失认(simultanagnosia)和知觉范畴障碍。狭义知觉性失认患者的基本视觉功能如视野、视敏度是正常的，但不能正确地辨识简单的视觉形态。同时性失认指的是患者在视野没有缺损的情况下，可以辨识一幅画面或一个情景的个别物体或人物，但不能对整幅画和面或情景正确地认识。双侧顶枕部损伤导致背侧同时性失认，患者数一个以上物品有困难，比如，目睹横排的一行硬币不能计数，某患者看窗外的马路，说“我知道那里不止一辆车，但我在同一时间内只能看到一辆车”。左颞枕叶结合部损伤导致腹侧同时性失认，同样不能辨识一个以上的物品，与背侧同时性失认不同之处是计数作业能完成、对自然环境的处理损害较少。知觉范畴障碍指患者不能正确地匹配从不同视角呈现的二维或三维物体，大部分是单侧右后半球病变所致。用Marr的计算理论分析，这些患者未能形成独立于观察点的描述，从而使非传统或非典型的观察点呈现的物体识别或匹配受损。总而言之，复杂的视觉能力由可分离的信息加工流组成，这些信息流包括物体形态、颜色、明暗、大小、移动、空间位置及整合，知觉性视觉失认是这些信息流中的某一个或几个受损所致。

针对知觉性视觉失认的评估有3个常用的测验：Gollin图形测验(由一系列逐渐变得完整的素描图组成)、不完全字母测验和不常见视角图片测

验。知觉性失认患者在这些测验的表现比正常人差。不常见视角图片是从一个很不常见的角度拍摄的照片，如茶杯的常见视角是从侧面拍摄的，不常见视角从正上方拍摄，只能看到杯子的底部和手柄。用 Marr 的理论解释，患者不能将不常见视角获得的信息转换为三维模型表征。进一步分析发现患者不能处理透视缩短的照片，而对缺乏显著特征的照片的识别则没有问题。

(2) 联络性视觉失认：患者对所见物体不能命名，也常常是不能表达其语义知识和功能特征、不能对其分类和归类。同一物品的不同呈现方式，如照片和实物的匹配不能是一个典型的例子。但保存复制和匹配以相同方式呈现的物体的能力，知觉相对正常，比如，能对知觉特征相似的目标进行区分。针对面孔、颜色和书写材料的识别障碍分别叫颜面失认(prosopagnosia)、颜色失认和单纯性失读症(pure alexia)，物体失认的患者未必合并这 3 种失认。它们可以单独出现，因为其视觉加工和识别依赖于不同的脑结构，但因为这些结构比较接近，在大的双侧大脑后皮质血管性意外时常常同时受累而出现各种不同的失认表现。

颜面失认的患者不能识别熟悉的面孔，但一般还拥有一些关于这些颜面的内隐知识，如该颜面是否著名等。颜面失认是因为特异性颜面加工机制损伤而不是精细辨别能力缺乏造成的。Farah 提出的双加工模型对整体分析加工与局部分析加工进行区分，认为颜面识别涉及整体分析过程，阅读涉及局部分析过程，而物体识别两者皆有。

范畴特异性命名障碍是联络性失认的一种特殊类型，文献报道的选择性受损的范畴包括颜面、动物、植物和食品等。大部分患者表现出对生物类比非生物类更容易出现命名障碍，可能是因为前者的视觉形态更复杂、彼此更相似、有更多的特殊名称的缘故，也可能是因为生物类语义记忆结构的局灶性破坏。视觉失语症也是一种特殊的命名障碍，患者不能命名视觉通道中呈现的物体，但可以通过其他感觉通道完成对同一物体的命名。例如，呈现一只靴子，患者可以准确地做出穿靴子的动作，但嘴里却说“帽子”，但听觉辨认和用手触摸靴子没有命名错误。患者表现为局限于视觉的物体和图像的语义性命名障碍，没有知觉障碍，是联络性视觉失认的一种类型。

3. 其他失认症　除了视觉失认，其他感觉通道的失认包括听觉失认、体感失认和触觉失认。听觉失认包括纯词聋、非言语声音失认、感受性失音乐

症和听觉情感性失认。纯词聋指识别音乐和非言语声音保持而听词能力被破坏;非言语声音失认指非言语声音识别被破坏而理解词语能力保存;感受性失音乐症指欣赏音乐的各种特点的能力受损;听觉情感性失认指对语言内容理解正常而对表达感情的吟咏的语言理解受损。体感失认包括疾病感缺失、偏侧躯体失认、偏侧疼痛失认、自体部位失认、异处感觉(指不能对皮肤接受的刺激做出正确的定位)和动觉性幻觉(自觉一侧肢体发生了长短粗细等变化)。

第三节　注意的相关脑区与偏侧忽视

一、注意的相关脑区

注意是心理活动对一定事物的指向和集中。注意依赖于许多脑区域的复杂的相互作用。控制注意的神经网络主要有6个区域:脑干网状激活系统、中脑上丘、丘脑后结节、前扣带回、后顶叶和额叶皮质。每个脑区域在某种特定的注意功能中起到突出的作用。

1. 脑干网状激活系统(reticular activating system, RAS)　负责警觉和注意的唤醒。细胞体位于脑干,与大部分大脑皮质有弥散联络,使大脑保持在恒定的注意状态。注意维持指在一段不间断的时间内集中注意力完成任务的能力,RAS很可能有助于注意的维持。双侧RAS或弥漫性RAS损伤导致昏迷。

2. 中脑上丘　要灵活分配注意,就需要从一个位置转移到另一个位置的能力。通过控制眼睛运动从视觉物品或位置转移注意的重要部位是上丘,它负责外来刺激快速进入视窝。尽管眼睛的移动和注意的焦点并不同步,眼睛的位置总是跟随在注意集中点之后。外周的物品或位置进入视觉中枢的过程伴随着眼睛飞快地跳跃式扫视。快速扫视是约120 ms的对外周新的视觉刺激作出反射性触发反应;而规则扫视是在意识控制下200～300 ms的反应。对猴子的研究表明,上丘在快速扫视中起关键作用,上丘受损后快速扫视消失而规则扫视不受影响,后者只有在额叶视区受损后被破坏。人类的核上性麻痹是以基底节的部分退行性变和上丘的特殊退行性变为特征。观察这些患者的日常行为可以了解上丘在注意控制中的作用:与人交谈时没有眼睛接触、

举箸就餐时不看碗碟。实验发现这些患者的注意转移有特异的问题。下丘在听觉通道的信息的注意转移中起着同样的作用。

3. 丘脑后结节　丘脑处于几乎所有感觉信息的十字路口，丘脑后结节在选择性注意中起重要作用。注意选择，指在一段时间内将精力集中在某一相关刺激或处理，而忽略不相关或分心刺激。在一项人类正电子发射计算机断层显像(positron emission tomography, PET)研究中，当呈现8个字母围绕着字母O的图像与单独呈现的O的图像相比，发现只有丘脑后结节部位的激活增加，原始视觉皮质区域并没有出现激活。从大量字母中识别字母O需要集中注意力过滤信息，这说明是任务的过滤而非其视觉复杂性使丘脑后结节部位被激活。

4. 前扣带回　前扣带回被认为是皮质和皮质下结构的分界面，整合警觉、唤醒、指向和过滤功能，结合当时情绪或动机选择恰当的反应。设想你站立在咖啡店外面的大街中间，突然注意到你周边有快速移动的车辆，你必须处理这些车辆的情绪意义(如紧张、惊奇)并做出反应(如离开)。这时，是扣带回赋予此景的情绪信息。功能影像学PET研究表明阅读Stroop色词测验的卡片C时激活在前扣带回。从大量的相关词(如斧头等)中选择一个词(如树木)时，也可以观察到扣带回激活增加。

5. 后顶叶　顶叶在注意的空间方面和注意资源的分配方面起重要作用。顶叶受损将导致偏侧忽视，患者对一侧空间往往是左侧空间的物品不注意。顶叶涉及注意资源的分配，因为出现在个体有意识注意任务中的特殊信息之时的P 300波幅在顶叶最明显。例如，在双重任务中，要求注意分割(指在一段时间内将注意力集中在一个以上相关刺激或处理)，受试者既可以给予相等的注意(任务A是50%，任务B也是50%)，也可以给予不等的注意(如任务A是20%，任务B是80%，或者反过来)，执行该任务时有意分配的注意力愈多，则顶叶区的P 300波幅愈大，这个P 300有时也被称为P3b。

6. 额叶皮质　额叶皮质对独特运动反应的选择和在充实服务于目标或计划的注意资源中起重要作用。额叶病变也可以观察到忽视，但不同于后顶叶病变导致感觉信息忽视，额叶病变是对忽视侧空间的运动性移动不能。额叶还驾驭注意指向的眼睛凝视，与眼运动的有意识控制和反射性眼移动的抑制有关。在P 300研究中，出现新的或不可预知的刺激试图引起受试者注意时，额叶中央区的P3a波幅下降，即额叶与新刺激的识别有关。

上述部位构成注意的神经网络，RAS 的主要作用是维持警觉和唤醒，扣带皮质是赋予信息以动机意义，后顶区提供感觉地图，额区提供通过探索、扫描、抵达和注视来转移注意的运动程序。但每个脑结构的特殊作用不是绝对的，比如，额叶是注意网络的一部分，同时也是其他许多认知功能的发源地，严重的注意障碍可能涉及 2 个以上脑部位。Posner 提出视觉选择性注意的三阶段理论模型是：注意从一个给定的视觉位置脱离，由顶叶控制（偏侧忽视被认为是一种注意脱离障碍）；注意从一个目标移动到一个新目标，由上丘控制；注意接近或锁定一个新目标，由丘脑后结节控制。

二、偏侧忽视

1. 定义　偏侧忽视（hemineglect）是突出的注意障碍之一，患者临摹图画时只画一侧（通常是左侧）的空间内容、阅读时忽略一侧的文字、书写时只在纸张的一侧、列算式时忽略一侧数字、吃饭时只吃盘子一侧的食品，严重者认为左侧身体不属于自己，极端者甚至认为这些肢体是他人的（称为躯体错乱，somatoparaphrenia）等。感觉对消（sensor extinction）现象是感觉性偏侧忽视之一，其表现为：在患者左侧呈现一支钢笔，患者能正确识别，钢笔放在右侧也能正确识别，而在左右两侧同时呈现钢笔时，患者仅能看到右侧的一支。顶叶缘上回病变易致偏侧忽视，右侧病灶比左侧病灶更常见、更严重。额叶、基底节和丘脑损伤者也有偏侧忽视报道。忽视症状通常在起病数周至数月后缓解，但一般不会完全消失。遗留症状的检查经常采用双侧同时刺激技术，分别在视觉、听觉和触觉通道同时呈现相似的刺激，患者会忽视一侧的刺激。

2. 临床表现　偏侧忽视表现为一侧景物看不见，尽管如此，但它不是左侧视野感觉加工损害的结果。忽视患者以身体中线为坐标，其左侧所有视觉材料的加工缺失。偏盲患者身体不动、只要转个头，注视左侧视野，就能看到原来看不到的物体，而忽视患者依然看不到。视野损害可以通过恰当的搜索策略作出弥补，而忽视患者不能。视野损害的病灶通常在枕叶，而忽视的病灶通常在顶叶，但如果病灶从顶叶延长到枕叶，忽视患者可以合并左侧偏盲。

3. 相关测验　偏侧忽视的检测有无意义结构模仿测验，字母、数字和符号划消测验，线条等分测验，两点辨别测验，自行画图、物品图形（如树、花）模仿测验，连线测验 A 与 B，文章朗读和文字抄写等，也有不同材料混合在一起

要求划消的测验，如 Caplan 字母与符号混合划消测验和 Halligan 等单词与星星混合划消测验。

在线条等分测验中，要求将一根线条分为相等两段，左侧忽视患者常常将中点放在右 1/4 处，左边长度占 3/4。然而，如果在等分之前要求患者注意线条左缘的某个明显的标记，则患者能正确地指出中点。这说明忽视是一种注意导向障碍。Mesulam 报告一例忽视患者在完成字母划消测验时，要求划掉图中所有的 A，患者表现为左侧忽视，当告诉患者每划掉一个 A 可以得到一元钱时，左侧忽视减轻了，这说明扣带皮质发挥作用改善了患者的注意动机，忽视的严重性可由注意力节制。患者对自己的偏侧忽视没有自知力，对他们来说，左侧空间是不存在的。这就像是脑后的空间，平时，你聚精会神地目视前方，对背后的世界浑不在意。只有在每回首一次可得一元硬币的奖励或黑暗的街道上有脚步声紧跟而来，你才会频频回头。总之，忽视是注意缺损的结果。

作者在检测 2 例右顶叶梗死患者的汉字阅读中发现左右结构的汉字容易出现偏侧忽视（如睛→青、腔→空、颠→页、勋→力），而上下结构（如尘、岩）与封闭结构（如园、困）的汉字罕见发生偏侧忽视。患者将“树”读作“寸”而不是“对”、“拗”读作“力”而不是“幼”、“谢”读作“寸”而不是“射”，字的左边和中间部分都被忽视了，而“纵”读作“从”而不是“人”、“诩”读作“羽”而不是“习”、“讹”读作“化”而不是“匕”等，字左边部分都被忽视，而中间部分没有被忽视，这说明患者的注意空间受到内隐记忆的影响。

为什么偏侧忽视易发在右半球损伤患者？不注意、肢体运动不能和指向运动不能（directional akinesia）也是右侧病灶比左侧病灶表现更常见。应用视觉搜索测验，有左半球病灶患者的平均遗漏项目左侧 1 个、右侧 2 个，而有右半球病灶患者的平均遗漏项目左侧 17 个、右侧 8 个。这说明右半球承担了更多的总体注意和觉醒功能。这些现象的一个解释是推测右半身活动受双侧半球支配，而左半身活动受右半球单侧支配，故右半球损伤患者易发生左半身偏侧忽视。另一个解释来自脑电图（EEG）的研究，左侧和右侧的视觉刺激都会导致右顶叶波的去同步化，而左顶叶的去同步化只在右视野受刺激时才表现明显，这说明右顶叶比左顶叶在警觉方面更占优势。

4. *治疗* 对于忽视患者，过街、骑车等简单的日常生活都会显得危险重重。冷热水刺激治疗可以减轻忽视症状。20 ml 7℃以上、体温以下的水在 15 s 内注入左外耳道，因为水顺着耳道前庭系统流动，患者不能保持平衡。冷

水送入使眼睛向刺激的对侧移动，而温水的导入使眼睛向刺激的同侧移动。遗憾的是，该技术会使患者出现眩晕和恶心等不良反应，限制了它的应用。可能的治疗机制是前庭刺激使颞顶叶区域血流量增加，从而暂时缓解各种忽视症状。

此外，视觉运动反馈、提高空间注意警觉性、左侧视野强化训练、戴右倾10°的棱镜等方法均可以改善忽视症状。

第四节　语言和语言障碍

一、语言成分的定义

语言是利用符号进行交流的工具。语言三要素为语音、语义和语法。音素(phoneme)是最小的可分离的声音单元，单独出现时毫无意义，按语言的规则把音素组合起来便得到了词素(morpheme)，即意义的基本单元。音韵(phonology)：通过研究不同声音的舌位、气流、口腔变化的模式来描述语言声音。字位、字素(grapheme)：意义的基本单元，语言文字系统的最小区别性单位。词位、词素(morpheme)：指语法的最小区别性单位，词是由一个或多个词素组成。

复述表示声音模仿或鹦鹉学舌。音素分析、言语发音的"内部联系"过程和适当的听觉语言记忆是复述的基本要素。复述能力的影响因素：音位结构、词汇级别、材料意义、字词长度、句法形式、可预测性、语法种类。

错语(paraphasias)包括言语错语(verbal paraphasias)、语义错语(semantic paraphasias)和音素错语(phonemic paraphasias)。言语错语指一个词语不正确地代替另一个词语，当这种代替的词语意义相似时(如以谷仓代替房子)，这种错误被称为语义错语，用声音相似的词取代目标词，被称为音素错语。

二、语言障碍的分类和表现

1. *失语症*　失语症可分为Broca失语、Wernicke失语、传导性失语、跨皮质运动性失语、跨皮质感觉性失语和完全性失语。各种失语症的鉴别见表1-4-1。

表 1-4-1 失语症分类与鉴别

失语症的类型	自发言语	错 语	理解	复 述	命名
Broca 失语	非流畅	不常见	好	差	差
Wernicke 失语	流畅	常见(口头上)	差	差	差
传导性失语	流畅	常见(文字上)	好	差	差
跨皮质运动性失语	非流畅	不常见	好	好(模仿言语)	差
跨皮质感觉性失语	流畅	常见	差	好(模仿言语)	差
全面性失语	非流畅	各种各样	差	差	差

(1) Broca 失语：构音肌并不瘫痪，丧失言语表达能力或仅能说出个别单词，口头表达的异常有 2 个特点，即在音素和音节水平发生障碍；语法类型减少和简单化。① Broca 失语表现为语音性和音位性障碍。音位性障碍类型：声音替代错误(一个声音节段被另一个音所取代)；简化错误(声音节段或音节被省略)；附加错误(增加了一个额外的声音节段或音节)；语言环境性错误(出现一个特别的音节，该音节不能通过周围的音位学的上下文的影响来解释)。② 听语理解：一般谈话正常，但不能执行 2 个以上的系列指令；抽象的无图片的项目、包含在句法形式中的信息的理解和阅读也有缺陷。③ 复述：严重受损，有的可复述 1 个或 2 个字词。④ 阅读(书写文字理解)：当句子的意义能被上下文及词语顺序清楚地传达时能理解，当句子的意义取决于相关的小的语法性字词时理解有困难；不能读出个别字母(字母失读症)，在日本患者表现为能读表意性汉字而读表音性假名的能力大大减低。⑤ 命名：很差，看图命名能力常略高于自发言语中生成名称的能力，原因为发音或记忆障碍。⑥ 书写：书写能力与说话能力平行，特点为拼音错误、字母遗漏、持续言语、句子语法缺失。抄写和听写好于自发书写。⑦ 运用：多见无瘫痪的左上肢观念运动性失用症和口面失用症。后者与左额叶盖部损害有关，表现为不能执行要求涉及面部和分音肌的动作，如吹气、弹舌、咳嗽等动作，自动性运动保存，如给患者一支正在点燃的火柴时他会吹灭。注意区别失用和听语理解缺陷。

(2) Wernicke 失语：① 自发言语流畅，语调正常，但用字错误，他人不能听懂，表现为音素性错语或语义性错语或两者同时存在。说话时：缺少实体词、对错语缺少自我改正的企图。音素性错语＋语义性错语＝新语性杂乱语(neologistic jargon)。音素性错语(在汉语语境中，这种情况少见，主要表现为语义性错语)(音素代替：如 team→[kimz]、省略 green→[gin]、添加 see→[sti]和换位 degree→[gadriz])；语义性错语：同范畴词代替 猫→狗 反义

词代替　好→坏　关联词代替　烟→火柴　总称代替具体　笔→铅笔。② 听语理解：严重受损。由音素编码水平加工过程缺陷引起(经皮质感觉性失语为语义水平缺陷)。依据：患者能匹配听觉呈现的词语与相应的图片。③ 书写文字(阅读)的理解能力：与听语理解受损程度平行。④ 书写：自发书写好于听写(可能有词聋)。⑤ 复述：不能。

Wernicke 失语的分类：Hecaen(1969)将 Wernicke 失语分为 3 个亚型：音素辨别缺陷为主型；语义理解为主型；注意力缺陷为主型。Kertesz(1983)则分为 4 个亚型：原始的 Wernicke 失语(说话流畅，伴有错语，但可被理解)；语义性杂乱语失语(说话大部分由语义性错语组成，不被理解)；新语性杂乱语失语(说话大部分由新语组成，不被理解，疾病感缺失)；纯词聋。

(3) 传导性失语：① 言语流畅，用字发音不准(音素性错语)。对言语和书写文字(阅读)的理解能力相对保存，患者能自我改正。② 不能复述，原因：音素解码异常(不能复述相对不常见的多音节词句)或短时记忆受损(记忆广度缺陷)。③ 有失命名症，患者强调知道正确的名称，但不能正确地说出。④ 病变：优势半球侧裂后部，主要累及缘上回皮质和皮质下白质，很少或没有神经系统异常体征。

(4) 完全性失语：表达和理解言语能力都丧失，但可保留执行体轴性命令，有的经数月或数年演变为慢性 Broca 失语。常见于左侧大脑中动脉主干的闭塞。

(5) 纯词哑：字词发音能力孤立的丧失而无书写能力、理解口头和书面言语能力的丧失。最常见于 Broca 失语的后遗症。与 Broca 失语不同点：即使在急性期也可用书写来准确回答问题并表达他的愿望和要求。复述严重受损，病灶：优势半球感觉-运动皮质发音区域，或 Broca 区，或该两区与它们传出通路之间的皮质下连接。

(6) 纯词聋：口语理解完全紊乱，患者说、读、写及非词语的听觉刺激处理并不受损。患者报告听到声音但没有字词，像听外国人说话。不能复述听到的言语。病变：左侧听放射和起源于对侧听觉区域的胼胝体纤维，接近 Wernicke 区(使听觉信息脱离)而又不破坏 Wernicke 区。最常见于双侧颞叶血管性病变后(两侧病变的时间可间隔数月或数年)。

(7) 无复述障碍的失语症：其共同特征是：保留复述能力(Wernicke 用"经皮质"一词指复述良好)，损害在侧裂区言语核心区域的外围。① 经皮质运动性失语：又名动力性失语。表现与 Broca 失语相同，自发言语的信息量和

复杂性明显减少。但复述好。如描述天气，患者只会说“好”，没有更多言语，但能复述“今天外面天气很好，是温暖的充满阳光的日子”。病变：Broca 区周围，或 Broca 区的前上方，或额叶运动前区。② 经皮质感觉性失语：听语理解在声音到意义的连接水平有缺陷。音素的听觉性和运动性加工无损害，故患者能复述他不懂的词句。自发性言语流畅，空洞。命名有明显障碍。可以是 Alzheimer 病早中期表现的一部分。病变：Wernicke 区后部，或左颞顶枕结合部的后和下部连接。③ 混合性经皮质失语：除了复述部分保留，所有言语功能均明显受损。病变：左大脑半球分水岭区，累及额、颞、顶叶大片区域。侧裂周围言语核心未受损。

(8) 失命名性失语：又称“命名性失语”或“健忘性失语”。口头与书面表达出现找词困难，而口头与书面言语的理解、朗读、复述、书写未受损。自发言语流畅，空洞，缺少实质词。病变：颞叶或颞顶区。分类：词产生失命名、词选择失命名、语义性失命名、范畴特异性失命名和感觉特异性失命名。① 词产生失命名：患者知道正确字词，但不能产生，是 Broca 失语或传导性失语的一部分；② 词选择失命名：患者承认不能回忆起恰当的字词，音素和语义暗示很少有效，但自称字词就在舌尖上，能从检查者提供的一组字词中快速地挑选出正确的字词；病变：颞枕结合处下部。③ 语义性失命名：检查者提供的某些字词，患者没有能力接受。病变：颞-顶结合部。④ 范畴特异性失命名：患者不能命名属于某一特异范畴的项目，如颜色，或身体部位，而其他范畴的项目能充分命名。最常见的是颜色失命名。病变：颞顶区。⑤ 感觉特异性失命名：包括视觉性失命名(大脑半球内视觉-言语连接障碍)、听觉性失命名和触觉性失命名。

(9) 基底节性失语：局限于纹状体的小病灶不产生言语障碍。病灶较大，可产生言语障碍。患者说话含糊不清，发音不准，但不偏离原来的音位，也不影响对其语义的理解。有些患者说话时，词与词之间缺乏连贯性。失语症的特点是口语表达中自发说话既非典型的流利型，也不是非流利型；如病灶靠后则倾向流利型。复述相对好，有命名障碍。对包含语法结构的句子有理解障碍，对文字理解的特点与听理解相似。

(10) 丘脑性失语：说话声低，音量小，有时甚至似耳语，但发音尚清晰。自发说话语量少。患者说话中有错语，以词义错语为主。复述相对好，对文字的理解有障碍，其特点与听理解相似。记忆缺陷可能是丘脑背正中核(DM)和板内区中的乳头体丘脑束的单独或同时发生病变所致；动机的缺陷是由于损

害了背正中核和它的额叶前部投射；言语表达缺陷是由于破坏了腹外侧核或它的额叶投射，而注意/意向缺陷是由于损伤了板内核的警觉通路或网状核或两者均受损所致。但是语义性缺陷可能在背侧丘脑前核或腹外侧核。背侧丘脑内侧梗死有明显认知缺陷，但语言保留。背侧丘脑后外侧梗死不产生神经心理学改变；而背侧丘脑外侧梗死时仅半数有言语障碍，但无失语，亦有轻的智能、视感知和结构改变。

(11) 交叉性失语：右利者由于右半球病变所致失语。原因：解剖畸形，缺乏皮质脊髓束交叉，已排除此种解释；遗传因素；偏侧化不完善；早期功能改组；偏侧化程度的皮质-皮质下分离，皮质下结构交叉性失语的发生率比较高。

2. 失读症

(1) 伴失语的失读症：① Broca 失语：书写缓慢而费力，因字母遗漏、颠倒而字词变形。打字不能改善(故不是利手瘫痪或肢体失用所致)。若书面言语相对保留提示病变局限、预后良好。② Wernicke 失语：颞叶前方病变，阅读理解相对保留；颞叶后方(颞顶叶)病变，阅读理解受损。③ 传导性失语、经皮质失语：书写与说话能力异常相平行。

(2) 不伴失语的失读伴失写：说话能力或听觉理解无任何障碍，只损及阅读和书写。病变：角回小梗死。视觉(印刷字体)、触觉(手指探索)和听觉(拼读)的缺损平行。

(3) 纯失读(失读不伴失写)：左大脑后动脉梗死导致左枕叶内侧、胼胝体压部病变，左顶叶未受累，书写和高声拼读保留。记叙性书写不费力，但不能阅读自己书写的内容。神经缺损：右偏盲、右偏身感觉丧失(累及丘脑)、词语记忆障碍(累及颞叶内侧)。

3. 语言障碍的文化背景差异

(1) 西语语言障碍的基本要点：① 口语流畅性、复述、听理解是区分不同类型失语的主要依据。Albert(1981)和 Prins(1978)指出，不到 20%～30%的失语患者会符合特异性失语综合征中的一种。② 各失语综合征划分的依据是言语行为障碍的模式，而非解剖定位或病理生理过程。③ 卒中急性期神经功能联系障碍(diaschisis)可掩盖缺陷的局灶性特征。④ 脑不是微型脑简单的聚集；脑区间的连接并不是汇合到一个执行中枢，而是采取关键脑区之间平衡对话的形式。

(2) 汉语语言障碍的争议：基于临床资料，汉语语言的脑结构定位研究有两种基本观点：一种认为汉语语言的脑结构定位与欧美国家的研究结果相

同，汉语与拉丁语的脑内接受、编码、保持和提取模式相似；另一种观点强调表意与表音语言的差异，认为汉人左脑卒中后失语的发生率较西方人低，汉人右脑有语言功能比例较西方人高，汉人脑卒中后失写症、失读症和空间忽视的发生率较西方国家低，纯感觉性失语和命名性失语极少，运动性失语最多。

香港中文大学心理学系陈瑞燕教授应用汉语认知激活范式（汉语成语填空测验、词语的语义和词形特征编码测验）和 fMRI 技术在香港英、汉双语中国人受试中的检测结果不同于上述两种观点，部分如汉字词形特征编码激活部位（左颞叶后下皮质）与表音语言的相似，部分如汉字编码与汉语成语填空的激活部位（左颞叶和左右两侧顶叶）与表音语言的不同，但还没有仅以汉语为母语的受试的检测结果。

为什么语言优势半球在左半球？左、右半球在语言加工中具有不同的作用是人类脑结构组织中最基本的特征之一。因为发音和手的肌肉的精细运动控制在左半球；聋的歌唱家的语言中枢亦在左半球。

右半球在语言加工中的作用：左半球常常被称为语言半球，但右半球并不是在语言加工中毫无作为。右半球在韵律和推论及叙事方面起重要作用。韵律（prosody）指语调模式或说出的声音的外套。叙事（narrative）是构造或理解故事线索的能力，而推论（inference）指"填写空白"的能力或根据隐晦的叙述材料作出假设的能力。当我们在说叙述句的时候，说话声音的高度通常是变低的；相反，在说疑问句的时候，说话声音的语调模式通常是变高的，因此，韵律线索有助于准确理解句子。在有些情况下，语调模式甚至是区分两句模棱两可的句子的唯一线索，比如，句子"他又受伤了？"表示说话者正在了解他是否再次受伤，而"他又受伤了！" 表示说话者已经知道他受了两次伤。文字一样，口气的侧重点不同，意义就截然不同。右半球与 Broca 区域同源部分在语言输出时韵律的产生起重要作用，而且，右半球与 Wernicke 区域同源部分在韵律的理解中起重要作用。作者曾经报道单纯右颞叶萎缩的患者没有失语症表现，但是叙事和推论能力明显受损。

4. 语言障碍的功能影像学研究　神经心理学经典的"病灶-缺损模型"，将失语症分为 Broca 失语、Wernicke 失语、传导性失语、经皮质感觉性失语、经皮质运动性失语、伴失语的失读（左角回受损）和不伴失语的失读（左枕叶和胼胝体压部受损）等。该模型的缺点有：病理学病灶很少与功能同种的神经解剖学系统相符；语言理解、发音和语法的神经基础是分布式网络结构，只有相对特异的定位点，患者采用代偿措施可以克服功能缺损，从而使病灶-缺损关系

不能建立；由于损害联络组织而使未受损的区域功能失调；如此等等，功能神经影像学研究可以弥补这些缺陷。

功能影像学（如 fMRI 和 PET）技术是受试者在独特神经心理测验条件下测量其局部脑血流量（rCBF）和脑葡萄糖代谢率（CMRglc）并揭示两者间的关系。认知激活的功能影像学研究使参与认知加工的神经活动历历在目，空间分辨率高。将激活条件的信号图像减去控制条件的信号图像就可以得到任务刺激作用的脑功能区。有关失语症患者的功能神经影像学实验对语言的贡献已经有了相当多的文献积累，本节介绍一些观点，是与传统神经心理测验不同的在线测验激活脑区域的研究成果。

（1）Broca 区损害对其远隔非损害区域的影响：知道 Broca 区损害导致语言表达障碍已经一个多世纪了，但这种损害对未受损区域有何影响是近几年才受到重视。Vandenberghe 等对 4 例左后下额叶皮质（属于典型的 Broca 区）受损的失语症患者进行 fMRI 检查，每例患者的语言输出受损，但对简单语句和书面语言的理解保留。在 fMRI 研究中，患者能够正常完成简单的非言语测验。视觉呈现一串字母（比如由 b，d，f，h，k，l 或 t 组成）、单词（比如 corn，space）和字母串（比如 svrn，mprzn），发现相对于字母串，识别单词出现的激活是内隐单词处理（implicit word processing）起作用。内隐单词处理是一种由 Stroop 效应揭示的心理现象。功能影像学研究表明内隐单词处理能力是包括 Broca 区的左额叶的功能。在 15 例正常对照组中，相对于字母串，识别单词出现的激活发生在右小脑、左后下额叶、左下顶叶、左后中颞叶、左后下颞叶和额叶上内侧。每个卒中患者呈现左后中颞叶皮质的正常激活，该部位与语义加工有关。然而，没有一个患者出现左后下额叶皮质（这是卒中受损区域）激活，也没有左后下颞叶皮质（这是卒中未受损区域）或正常对照组没有被激活的其他区域被激活。

简言之，左后下颞叶皮质作为卒中未受损区域在正常对照组中被激活而在卒中患者组没有被激活，左后下颞叶的反应依赖于左后下额叶皮质的传入。由此，功能影像学研究已经表明某个病灶对于远隔非损害区域的功能的影响，说明神经系统的整合特性。

（2）某个脑结构对于某个功能的必要性和充分性：Broca 区损害的患者言语输出受损、简单词汇的理解保存。以此推断，受损区域对于言语表达是必需的，但对于言语理解不是必需的。然而，根据“病灶-缺损模型”，该区域对于言语表达并不是必不可少的；相反，针对正常样本的 fMRI 检查揭示该区域对一

个言语表达任务尽管是需要的，但不是必需的——某些激活区域对于执行任务是多余的。因此，识别功能相关区域与病灶-缺损模型提供了特定区域对于该任务是否必需的信息。一个完成某任务必需的皮质区受损的患者，在完成该任务时，其能力可能依然保持，因为受损区的功能要么由不同的脑结构贯彻（如涉及对侧半球同系部分或认知重组），要么在常规结构影像学检查显示受损的区域仍维持其残余功能。所以，尽管在正常对照组完成该任务时被激活的区域受损了，患者在功能影像学检查时仍保持完成该任务的能力，该结构对于该功能就不是必不可少的。总之，必须结合正常和患者的功能影像学检查才能揭示某个脑结构对于某个功能的必要性和充分性。

语义决定任务运用的是单词和物品图片的语义相似判断，如 sofa 和 chair，哪个与 table 语义相似，lemon 与 apple，哪个与 orange 语义更接近。基线任务使用的是同一刺激的知觉判断（要求判断单词的大小，如 table 与 table，哪个与 table 一样大）。这个范式与左半球的额叶和颞叶区被激活有关，包括前下额叶、前和后下颞叶、颞中皮质、双侧后下顶叶皮质、前扣带回皮质和右小脑。第一个患者（缩写为 SW）能完成这个范式，尽管存在左额叶（包括左前下额叶）和颞顶叶皮质的广泛受损。功能影像学检查揭示 SW 完成语义决定任务时左颞叶和顶叶被激活，但他的左前下额叶皮质没有被激活，而所有的正常对照样本均能观察到该区域被激活了，SW 的右下额叶皮质也没有代偿性的激活。这说明正常对照样本完成语义决定任务时左前下额叶皮质被激活，对于 SW 完成同一任务并不是必需的。

患有语义性痴呆（semantic dementia, SD）的患者语义知识和命名进行性减退，而其他认知功能和语言功能相对保持正常，其解剖学基础为起于前颞极，随着病情进展向后扩散的双侧颞叶皮质萎缩。根据"病灶-缺损模型"，前颞叶区域是语义和命名损害的部位。应用上述语义决定任务观察 6 例轻度 SD 患者在 fMRI 检查时的表现，有 4 例按要求完成了扫描仪呈现的任务。结果发现激活涉及左下额叶、左颞顶叶、左前中颞叶皮质、前扣带回皮质和右小脑。与正常人不同的是患者的左后下颞叶皮质或右颞顶叶会合处没有被激活。患者的前颞叶皮质的激活比正常人更明显。因此，这项研究表明，损害的部位被激活了；与损害部位远隔的功能障碍；在损害的前下颞叶皮质和未受损的后下颞叶皮质之间的功能整合失调。患者的左后下颞叶皮质没有被激活说明完成该任务不需要该部位配合。在这个例子中，左下额叶皮质和左前中颞叶皮质出现代偿性的激活，提示存在认知或（和）神经重组。

以上研究表明，在语义相似判断任务中，SW 未能激活所有的正常对照组样本能激活的左前下额叶皮质，SD 患者未能激活左后下颞叶皮质。综合这两个结果，左前下额叶皮质和左后下颞叶皮质被激活对于完成该任务不是必不可少的。而左中颞叶、左前下颞叶、左后下顶叶皮质和右小脑在 3 组受试者（正常人、SW 和 SD 患者）中均被激活，表明这些脑结构对于语义功能的充分性。进一步的研究要选择病灶在这 4 个区域的患者。

（3）伴深层失读患者的右半球阅读是否属于神经系统副本：在正常大脑中，支配特定认知功能加工的神经结构可能是双重的。占优势的结构抑制了其副本。当占优势的结构受损时，较少主导的结构可能做出反应。因为失抑制后功能仍然保持，该结构受损所致认知损害也许从未被神经心理评估所识别。相比之下，功能影像学检查能揭示认知功能多重的、充分的脑结构基础。例如，左半球受损后，右半球能否保持某些语言功能。Blasi 应用单词提取学习任务的研究支持在左半球受损后右半球有代偿作用。

在左半球广泛受损后，右半球保持某些语言功能的假设是通过对深层失读患者的研究提出来的。深层失读以失语与不能大声阅读最简单的假词（如 gam，dake）、缺乏具体指示物的功能词（如 of，that）和抽象词（如 process，involve）为特征。尽管不能运用拼写和发音相符的知识读音，患者阅读有具体指示物的单词（如物品名称）的能力相对保存。最有趣的现象是患者针对有些单词发出语义相关但拼法和音韵无关的单词，如将 spirit 读作 whisky，将 yacht 读作 ship，将 bush 读作 tree。很明显，在这些例子中，深层失读患者一定程度上理解这些单词但没有产生正确的读音。

针对这个症状有两种主要的解释。一种是深层失读来源于残余的左半球阅读系统，该系统丧失了对书面词（这些书面词没有意义方面的提示或参考）正常的发音能力；另一种解释是由于左半球广泛受损，患者依赖右半球阅读系统理解单词。功能影像学检查可以检验这些神经心理学假设。假如深层失读患者的阅读能力依赖于残余的左半球阅读系统，患者在阅读描写物品的单词的时候，左半球词语加工区应该被激活。另一方面，如果患者右半球翻译单词的意义、再由左半球输出言语，可以预期右半球（而不是左半球）的语义相关区域被激活，同时左半球的言语输出相关区域被激活。Price 等报道了 2 例深层失读患者（JG 和 CJ）借以检验这些预测。

两个患者均在左半球语义和音韵系统剩余区域被激活，这排除了深层失读患者纯粹由右半球加工的解释。然而，相对于正常人，患者右半球也提高了

活动性。在 1 例(JG)患者中,右半球的额盖和后下额叶皮质被激活,左半球相应的部位没有被激活。左半球的额盖和 Broca 区与发音及言语输出有关。因此,在左半球受损后,JG 的右半球能够发音。语义区激活发生在左半球,没有证据提示存在右半球语义系统副本。患者 CJ 右前下颞叶皮质和右额盖被激活。左前下颞叶皮质是语义加工相关区域,因此,对于患者 CJ,右半球也有语义加工作用。总之,深层失读的功能影像学检查已经表明右半球语言区存在过分活动,但激活的区域在 2 例患者中不一致。另外,由于 2 例患者是在脑梗死多年后接受研究的,难以判断激活反映的是失抑制还是潜在的长期改变。

(4) 失语症患者语言功能的恢复:患者如何恢复失去的功能是对脑损害患者进行影像学研究最关键的问题之一。功能影像学研究已经提出 3 个可能的恢复机制:① 梗死灶周围激活,环绕病灶的活组织能够调整以往由该区域更多细胞维持的功能。② 神经重组,当认知功能需要不同的神经结构实现其功能的时候可能出现神经重组。③ 认知重组,当患者为了弥补失去的功能,使用不同的认知形式(和不同的神经结构)完成同样的测验的时候,出现认知重组。

使用传统的 CT 和 MRI 扫描获得病灶的形态学指标不一定意味着功能的完全丧失。有时病灶很大,患者的恢复却令人惊讶,出人意料得好。相比之下,病灶的功能影像学指标能识别受损区域的功能保持状况。丧失功能的恢复要么是开始未被激活的组织的重新激活(如由于水肿消退),要么是活组织神经突触效力增加直到能够支持原来由更多细胞承担的功能。梗死灶周围活动在功能恢复上的重要性可能被严重低估了。这是因为以往绝大多数功能影像学研究是汇聚不同的患者的数据资料作为一个组别、分析比较患者组和正常组的差异。患者的梗死灶的大小和位置千差万别。对梗死灶周围激活的研究应该具体病例具体分析。

在 Warburton 的研究中,选择 6 例左颞顶叶区域有大梗死灶的失语症患者,于静态和单词生成任务激活情况下分别进行 6 次扫描。每例患者的资料单独分析,并与 9 例正常人的表现比较。在正常样本,单词生成任务一致地激活了左半球语言区的广泛区域,尤其是左前额和后颞叶皮质。一半的正常人在右半球的相应区域也被激活。所有 6 例患者的左前额区被激活,病灶最大的患者的左颞叶被激活了。右前额和颞叶皮质也有一半被激活。将患者作为一个组别的时候,不能识别左颞顶叶梗死灶周围的激活情况。总之,除了病灶周围的激活,恢复期失语症患者在单词线索提取相关的激活与正常对照组难

以鉴别。多个类似的纵向研究表明失语的恢复与梗死灶周围左半球语言区的重新激活有关。Leger 针对一例左半球受损导致言语输出障碍的失语症患者治疗前后分别做 fMRI 检查，发现语言治疗作用与 Broca 区及左缘上回的激活有关。

Weiller 针对失语症的功能影像学研究提示恢复发生在侧性转移后，即受损区域的功能转移到对侧同系部位。侧性转移来源于右半球语言区副本的失抑制或已经存在的语言区的反应增强。已经恢复的失语症患者的右半球语言区比正常对照组有更多激活。为了检验神经重组已经发生，患者和正常人的针对激活任务的认知过程应该相同。但有些认知功能如注意定势、内隐认知加工和操作水平还不能设计出认知激活模式。总之，失语症患者语言恢复的机制还不是很清楚。进一步的研究必须在认知改变和神经改变之间做出区分。尽管如此，功能影像学检查已经清楚地表明某些失语症患者的语言恢复是凭借受累皮质的残余功能，运用的是和正常人一样的神经系统。

综上所述，Broca 区损害对其远隔非损害区域（如左后下颞叶）的功能有影响。通过对左额颞顶叶皮质卒中患者、语义性痴呆患者和正常对照组样本在语义相似判断任务中被激活的区域比较，发现左前下额叶皮质和左后下颞叶皮质对于语义识别不是必需的，而左中颞叶、左前下颞叶、左后下顶叶皮质和右小脑在 3 组受试者中均被激活，表明这些脑结构对于语义功能的充分性。右半球阅读是否属于神经系统副本及失语症患者语言功能恢复的机制还没有一致的结论。

第五节　记忆的相关脑区和遗忘症

一、记忆的定义与分类

记忆是使储存于大脑内的信息复呈于意识中，保存和回忆以往经验的过程。从信息加工的角度可将记忆分为 3 个阶段：获得是通过学习在大脑留下记忆痕迹的过程；储存和巩固（编码）是记忆痕迹从开始时的不稳定状态逐渐转化为长期牢固并储存下来的过程；再现（提取）是将储存在脑内的记忆痕迹回忆出来的过程。

以时间间隔来划分，可将记忆分为即刻记忆、短时记忆和长时记忆。即刻

记忆又称感觉登记，是感觉器官获得的短暂保留的信息(持续 0.25～2 s)，完全依据它所具有的物理特性编码，视觉感觉记忆保留 0.25～1 s，听觉感觉记忆在 4 s 内，很容易被新的信息取代或自行消失，只有受到特别注意的材料才能转入短时记忆。短时记忆指在意识中暂时掌握的信息(1 min 内)，容量为 7 ±2 个组块。工作记忆被认为是短时记忆的一种类型，由 3 部分组成：中枢执行系统，是连接长时记忆的注意控制系统；发音环路，听觉材料被存储和复述；视觉空间映像：视觉短时储存。延迟记忆一般指保持 1 min 至 1 h 的信息。有的将它归为长时记忆，有的将它作为短时记忆与长时记忆之间的中间记忆。长时记忆指能够长期保存的信息，有的可以维持终生。

根据意识是否参与，可以将记忆分为外显记忆(又称有意记忆)和内隐记忆(又称无意记忆)。外显记忆可以进一步区分为语义记忆和情景记忆。语义记忆指与环绕我们的世界相关的概括和不变事实的外显记忆，以语义知识为基础。语义知识是以复杂的概念联系网络加以组织的，概念的属性是事物相互区别和归类的基础。情景记忆指个人体验的外显记忆，包括它们的时间、地点或空间背景及曾经的感受，提取时需要努力、缓慢、谨慎的记忆检索。内隐记忆包括知觉记忆和启动效应。

记忆不是单一的过程，记忆是有机协调组合在一起的一组系统，每一部分的记忆都有各自不同的功能，由不同的大脑组织支持。正常的记忆功能需要许多大脑组织协同作用、共襄其事。肯定的与记忆的相关脑区包括内侧颞叶(海马系统)、前额皮质、间脑和杏仁核等，有争议的相关脑区包括颞叶新皮质、基底节和小脑等。

二、记忆的相关脑区

1. 内侧颞叶(海马系统)　颞叶内侧区包括海马、内嗅皮质、嗅旁皮质和海马旁回。海马系统接收所有来自大脑新皮质处理器的汇合输入信息，接收关于人和物、时间和空间背景、情感和行为反应，以及构成学习经历的行为的信息。各个皮质加工区处理特定场景或事件的各种要素，如视觉、听觉、语言和空间信息，事件的不同属性和成分可以储存在不同的新皮质区，如情景经历的视觉部分记忆被储存在视觉加工区，语言部分记忆储存在语言区。海马系统和这些不同的新皮质系统之间都是相互连接的，以便海马系统与加工原始事件或场景有关的新皮质区域保持沟通。因此，对完整事件的记忆是分布式储存的。

外显记忆(尤其是情景记忆)的中枢神经结构是海马系统和相关结构。外

显记忆又称为陈述性记忆，是需要对以往事件或体验进行自觉追忆的记忆，需要对信息作有意识的重建。外显记忆主要取决于海马系统，还有以下证据：① 在 fMRI 研究中，受试者在执行一项需要多个信息流相结合的任务中，海马系统被激活。② 对海马系统的神经解剖学和生理学检查表明海马系统拥有支持这种关系连接所必需的解剖连接和神经机制。对于大脑处理的不同种类和形式的信息，海马系统从新皮质区接收信息输入，回馈投射至新皮质区。海马系统还扮演处理连接记忆机制的角色。它表现出一种特别强健的突触可塑性形式，叫做长时强化（long-term potentiation，LTP），特定通路的短暂激活可以产生持续数小时乃至数周的稳定的神经突触效力增强。这种可塑性形式是通过 NMDA 受体传播的，后者组成强有力的结合传感器，只有时间上很紧凑的信息输入汇合才能被激活。③ 海马系统的电生理研究发现神经元的活动代表着对环境中各种有意义的线索和物体关系的编码。例如，当大鼠积极探索周围环境时，海马神经元拥有位场（place fields）。当动物处在环境中的某一特定“位置”时，对应的细胞优先被唤起。神经元的表现和细胞唤起有关，这一点并不取决于众多的环境刺激中的哪一个，而是取决于刺激物之间的关系。

2. 前额皮质　工作记忆是短时记忆的一种类型，储存在背外侧前额皮质，由 3 部分组成：中枢筹划系统，是连接长时记忆的注意控制系统；发音环路，听觉材料被存储和复述；视觉空间映像：视觉短时储存。早先对于记忆的理论研究认为工作记忆和长时记忆处理信息的方式是按照一种严格的顺序：首先由工作记忆保存信息，然后，假如需要记住该信息的时间达到一定的长度，则转为长时记忆储存。记忆巩固则被认为是将不稳定的短时记忆储存转到更稳定的长时记忆储存的过程。目前认为工作记忆和长时记忆储存在不同的脑区。遗忘症只损害长期记忆，工作记忆却可保持。后者使得遗忘症患者可以像正常人一样理解情景、事件和语言，可以进行有条理的谈话。与遗忘症相反，在有些病例中，工作记忆出现选择性损害而长时记忆完好，这种现象形成了两种记忆系统的双重分离。对存在选择性工作记忆损害的患者进行研究，以及对在工作记忆功能中很关键的背外侧前额皮质区域进行解剖学、生理学和功能影像学研究使我们对工作记忆有了更完整的认识。工作记忆和长时记忆必须被看作是并行不悖的记忆系统，工作记忆系统将信息保持在活动状态以支持在线处理，长时记忆系统则创造持久的经历纪录以备后用。关于工作记忆缺陷最著名的一个例子是有关听-口语工作记忆损害的。患者不能大声复述和一字不差地重复就在片刻之前的发音，但是，能够保留和恢复口语输

入链中的信息，甚至可以学习单词表。他们在其他处理领域内的工作记忆（比如空间处理或者代数运算方面）却完好如初。有人还报道过存在视-语工作记忆缺陷的患者。因此，每一种这样的缺陷都和专门的处理区域相关，其他处理域的工作记忆完好。这意味存在着多个工作记忆容量，每一个都和大脑的特定的信息处理模块的运作紧密相连。前额皮质似乎还有不同的部分，虽然它们都和工作记忆有关，但每个部分都致力于某种特定的工作记忆类型。

策略记忆的储存在前额皮质。策略记忆有很多方式，如时间顺序的记忆：比如判断连续呈现的两个项目哪一项是最近出现过的；来源记忆：回忆特定信息的来源；自我定序指示任务：在许多项目中立刻指出一项，每次指出一个不同项目，各项目在队列中的位置则是经过处理的，要求说出每格空间位置项目。前额皮质损伤，影响到的不仅是工作记忆。比如外伤病灶造成的前额皮质损害，还可产生策略记忆的缺陷。虽然有此损害的受试者并不存在对刚才学习过的项目的再认困难，但是，如果超出这个范围，当项目回忆必须基于上下文背景时，要求他们重建当时的学习环境或深入推断项目信息，会有受损表现。衰老对额叶功能的影响比对大脑其他部位的功能影响大，即使老人没有神经系统损害，比起年轻人，也可有策略记忆缺陷表现。年迈显然对额叶有着异乎寻常的损害。

3. 间脑（乳头体和丘脑） Werniclce-Korsakoff 综合征是被研究得最多的遗忘症病因，因多年慢性酒精滥用及硫胺缺乏导致间脑中线结构损害所致。中线丘脑核损害可累及海马系统，因为丘脑和海马相互连接。也有许多患者损及额叶，除了遗忘症状，还有额叶功能障碍，如策略记忆受损。策略记忆损害并非遗忘症出现的记忆损害的必要部分。相反，只有同时存在额叶功能障碍，才会出现策略记忆方面的问题。还有一种观点认为该症是丘脑和额叶之间的联系中断所致。丘脑卒中是遗忘症的另一种常见病因。丘脑卒中最容易引起血管意外的是丘脑膝状体动脉（大脑后动脉的第 2 节段），其次是丘脑-下丘脑旁正中后动脉（大脑后动脉的开始节段基底交通动脉），通常呈双侧（蝴蝶结状梗死）。有记忆损害的丘脑卒中累及内髓板和乳头丘脑束。乳头丘脑束是 Papez 环路的组成部分。与颞叶内侧病灶导致遗忘相比，间脑病灶导致遗忘有更多的逻辑记忆、编码学习方面的损害，有更多的虚构现象。

4. 杏仁核 情绪记忆既不是外显记忆，也不属于内隐记忆。比如，惊恐障碍患者呈现显著的焦虑情绪，但不能识别出特异的病因。情绪记忆的关键结构是杏仁核，它与控制自主功能（如血压和心率）的系统、与下丘脑及激素分

泌均有联系。杏仁核破坏后情绪记忆被革除，但对外显和内隐记忆几乎没有影响。相反，严重痴呆患者已经说不出照片上人的姓名，但由于其杏仁核结构未被累及，还能根据自己平时好恶讲出喜欢或不喜欢照片上的人。

三、遗忘症

1. 遗忘症的病因　根据自然发展过程划分遗忘症的病因：① 突然起病的遗忘症，通常会逐渐但不完全恢复，如双侧或优势侧海马梗死、双侧或优势侧前内侧丘脑梗死、额叶底部梗死、蛛网膜下腔出血、间脑与颞叶内侧或额叶眶部的外伤。车祸致头部损伤往往导致颞叶内侧和额部损伤、心脏骤停、CO中毒和其他脑部缺氧状态（如高山症）致海马损伤，长期反复癫痫发作之后。② 突然起病、病程短暂的遗忘症，如颞叶癫痫、电抽搐治疗之后、短暂性全面性遗忘症。③ 亚急性起病伴有各种恢复程度的遗忘症（一般会遗留永久性损害），如 Wernicl ce-Korsakoff 综合征、单纯疱疹性脑炎、结核性或其他原因的脑膜炎。④ 慢性进展性遗忘状态，累及第三脑室底部和边缘皮质的脑瘤、退行性痴呆症、类肿瘤“边缘”脑炎。

2. 顺行性遗忘的特点　顺行性遗忘是选择性地损害长时记忆能力，而工作记忆则完好无损，即可以保留当时正在加工处理的少量信息。相应的，假如在材料获取和记忆测试之间间隔时间比较短，遗忘综合征患者可以有正常的表现，但假如间隔时间再长一点，他们就无法保持记忆了。数字广度测试提供了工作记忆的指标，扩展数字广度（即在数字广度基础上增加一个数字）则提示长时记忆，两种测试可以发现遗忘综合征患者长时记忆和工作记忆的分离。由于患者的记忆损害在显著迟延后就会“暴露”出来，他们无法持久保存信息。因此，他们无法学会和累积关于世界、关于自身的新的事实和数据。顺行性遗忘损害的只是长时记忆中的一个特殊部分，亦即对新事实和新事件的记忆受损。在大多数此类记忆测试中，要求受试者学习一组常用单词、面容、视觉物体等等，遗忘症患者在回忆或者辨认曾经学习过的项目时都存在显著的损害。顺行性遗忘并不累及熟练行为的获取和表达。

虽然在许多记忆测试中，遗忘症患者都表现出显著的能力受损。但在某些情况下，即使是最严重的遗忘症患者，其内隐记忆完好无损。内隐记忆又称为程序性记忆，指自动的、不需要有意识回忆的记忆现象，提出由于先前的经验或行为得以易化或改变，它包括启动效应、条件反射、习惯化和运动技能学习等。旋转追踪测试、谜塔推理测试、不完整绘图识别测试、镜像阅读测试、词

干完形测试及眼动监控技术等大量检测证实这种保留。以下仅列举词干完形测试说明其设计理念。

词干完形测试：向受试者提供一个词汇表供其学习。表中每个词的前 3 个字母是单词的“词干”。例如，学习词表中包括 motel 和 cyclone，它们的前 3 个字母也是单词 mother 和 cycle 的开头字母，而后两个单词则不在学习表中。经过一段间隔时间以后，用两种方法测试受试者对这 3 个字母词干的记忆，得到两个结果。一种测试方法是线索回忆，在这种测试中，遗忘症患者的表现比起对照组来要差得多。另一种测试方法是词干完形，要求受试者讲出他想到的第一个由该词干组成的单词。遗忘症患者的表现正常。因此，和对照组的受试者一样，遗忘症患者也表现出对前面信息的启动反应，因为他们也倾向于采用学习词表中的单词来完成词干组词。简言之，针对可以通过培训改善的能力和各种关于培训经历的记忆，遗忘症患者的表现是不同的，完好的是重复启动和技能学习，受损的是自由回忆和辨别。

顺行性遗忘者能够获取的信息依赖特定环境，即只能体现在特定的有来龙去脉的材料中的信息。在技能学习和重复启动测验中，出于不同需要进行的评定都是以同一方式进行的：每一次都是对原先环境的重复。患者在启动和技能学习测试中表现出来的完好的行为能力并不涉及其他测试内容。早先的学习过程即使再现，他们也不记得曾经见过。只有当材料再现，测试条件和原来的学习条件在关键方面都一模一样时，才能发现患者具有学习能力。

3. 逆行性遗忘的特点　不同的逆行性遗忘症病例其病程跨度存在显著差异：有些持续几十年，而有些只有几个月或几年。逆行性遗忘可以区分为两大类：长时逆行性遗忘，影响到几十年的记忆量；限时逆行性遗忘，只影响几个月或几年的记忆。逆行性遗忘综合征的时间跨度因其病因而有所不同。柯萨克夫病、阿尔茨海默病、帕金森病和亨廷顿病都会产生长时逆行性遗忘。相反，其他患者，最常见的是闭合性脑外伤患者，表现为限时逆行性遗忘。在病情稳定后，经过这种伤害的患者多数表现出非常短暂的逆行性遗忘。有人研究过 1 000 多例的连续病例，其中 80％以上的患者有过 1 周内的逆行性遗忘。

多数逆行性遗忘综合征病例，不论其病程长短，都表现出一种时间梯度，即遗忘综合征影响近期记忆，甚至于影响远期记忆。尽管柯萨克夫遗忘综合征患者的缺陷跨越测试的任何阶段，但其最近期的缺陷最严重，时间越远的记忆所受影响越小。有些时候逆行性遗忘是没有病程起伏的。阿尔茨海默病晚期、某些脑炎后患者，以及亨廷顿病患者损伤的区域可能正是长时记忆的储存

区，造成没有时间梯度的逆行性遗忘。有的临床医生在查房时常常将遗忘综合征患者不能讲出自己今天早饭吃了点什么或者他们现在住在哪里作为一种“短时”记忆缺陷，而将患者能够讲出自己的姓名、出生地等等说明其“长时”记忆完好无损。这是他们误解了“短时记忆”和“长时记忆”。心理学上的短时记忆指的是个体积极处理信息阶段时用的记忆（比如立即重复7个数字时采用的记忆）。正如我们前面已知，遗忘综合征患者可以记住信息的时间非常短；这就是说，他们有正常的短时记忆。这些记忆的时间段是以秒、分来计，而不是年、月、日。因此，最好采用近时记忆（recent memory）和远时记忆的概念来区分逆行性遗忘中丧失的不同时间段的长时记忆。

成年逆行性遗忘均不会永久性影响所有的长时记忆，因为患者早年学到的许多信息都保留了下来。不论逆行性遗忘多么广泛，除非患者又发生了别的广泛性新皮质区损伤，而该区域刚好是储存长时信息的所在，他都可以保留早年熟练掌握的信息。因为对所有人而言，这一部分信息都是知觉、运动和语言能力的基础。柯萨克夫病、缺氧和脑炎而引起的遗忘综合征患者，虽然其逆行性遗忘是广泛的，但其对世界的知识完好无损；保留语言、知觉和社会技能；一般智能正常。除非进行性痴呆晚期才会损害此类信息。这也说明海马系统及其相关构造并不是所有长时记忆的储存地。所有短期逆行遗忘综合征的患者其远时记忆完好无损，甚至长时逆行性遗忘患者都仍然可以保留早年获取的信息。相反，新皮质似乎才是长时记忆的永久储存地。

正如顺行性遗忘一样，在逆行性遗忘中，记忆支持的技能行为也完好无损。在顺行性遗忘中可以见到，记忆的直接测验（如自由回忆）受损而间接测验（如启动）中完好无损，这种记忆分离现象在逆行性遗忘中也有发现。

四、元记忆

1. 元记忆的概念　1976年Flavell首先提出“元记忆（metamemory）”的术语，用以指对记忆的认知——特别是对不同记忆任务和不同记忆条件下的记忆要求，以及在相应条件下可以改善记忆的策略的意识，是关于自己记忆能力和策略的知识与信念。Wellman的看法更明确而容易被人接受，他认为元记忆也是一个认知系统，包括对记忆任务、过程和策略的知识；对当前记忆内容和状态的自我意识，对自己记忆程度的判断和评估，知道是否记住了，自己知道什么或者不知道什么；对个人记忆能力全貌的自我评价，以及与记忆有关的情绪和情感状态的体验。

2. *元记忆的意义* 客体记忆(或称记忆,即有关信息获得、保持和提取的记忆)与元记忆可能依赖于不同的神经生物学机制,元记忆可能比客体记忆更多地依赖于前额叶的功能。Janowsky 等通过对额叶损伤患者的研究,直接得出了前额叶与知晓感判断(feeling-of-knowing,FOK)密切相关的研究证据。Schnyer 等对额叶损伤患者进行提取自信度判断(judgement-of-confidence,JOC)和 FOK 研究结果提示两者是可以分离的,JOC 相对保留完整而 FOK 明显受损下降,同时对 FOK 明显受损患者的影像学进行叠加,显示损伤主要集中在右侧前额叶腹内侧(ventromedial prefrontal cortex, VMPC)。Maril 等采用功能影像学(fMRI)研究提示前额叶出现了与 FOK 相关的激活。Kikyo 等的 fMRI 研究亦提示双侧额下回(inferior frontal cortex, IFC)、左侧额中回、前扣带回及辅助运动区和双侧尾状核均参与了 FOK 的加工。随 FOK 的提高,其激活信号呈线性增强,而其中双侧 IFC 在能成功回忆的过程中激活不明显,这提示 IFC 可能在元记忆监控系统中起着重要作用,IFC 的激活是 FOK 认知的基础。Chua 等最近通过 fMRI 证实海马、扣带回和其他边缘结构(以往称为 Papez 环路)是 JOC 的基础,从而提出元记忆有 2 个网络结构基础,即颞叶内侧(medial temporal lobe, MTL)和前额叶腹内侧(VMPC),前者可能是 JOC 的结构基础,后者是 FOK 的结构基础。

Jonker 等认为元记忆损害是痴呆的预测指标之一。元记忆损害的识别有助于进一步理解和判断脑老化的异质性,有助于了解和明确"认知障碍的就诊率为什么非常低""老人的记忆主诉往往不准确"等临床经常观察到的问题,有助于制订恰当的药物治疗方案、记忆改善与训练干预策略。

3. *元记忆的评估方法* 元记忆的评估方法有临床等级评价法、问卷调查法、心理测量法、现象描述法和综合法。研究文献主要集中在心理测量法,包括测验受试者主观报告、知情者评价、施测者客观测量评价的比较、使用结构性晤谈的行为观察记录、实验性元记忆范式的运用,后者包括任务难度判断(ease-of-learning, EOL)、学习程度的判断(judgement-of-learning, JOL)。JOC 和 FOK 研究中 FOK 被研究最多。Nelson 指出这几种元记忆的评估方法没有明显的相关性,它们分别监测调控记忆的获得、保持和提取等某个环节。FOK 的实验研究范式是在传统记忆研究方法回忆和再认的过程中间加入判断,即 RJR(recall-judgement-recognition)模式。FOK 是从记忆储存中完全提取信息失败时的情感体验,仍然能够提取与目标相关的部分信息,从而作出对将来是否能够再认出目标的判断。Nelson 等发现在对那些确定知道但不

能立刻回忆出的项目所花费的时间更长，提示 FOK 是一种对在记忆中有效回忆提取的复杂搜寻过程。FOK 是对记忆的提取阶段进行监测的一种中间状态，在对记忆绩效方面起着基本的调节作用。

4. 元记忆的影响因素　元记忆的影响因素非常多，Clare 于 2004 年提出元记忆的生物-心理-社会模型：① 生物学因素：认知损害的严重度和脑影像学特征，神经认知特征，如记忆、执行领域的知晓感。② 心理学因素：如应对风格和病前个性特征，尤其是基于自信度判断的 FOK 和 JOC，受人格倾向、行为模式和情绪状态影响。③ 社会因素：受试者与朋友、亲戚和保健专业人员的相互作用。

第六节　执行功能的相关脑区

在神经心理学理论体系中，执行(executive)功能是一个非常重要的概念，大部分神经心理学专著如 Lezak 的《神经心理评定》与 Banich 的《神经心理学》，将其列为单独一个章节予以论述。国内研究者将 executive 直译为“执行”，但是，在汉语里，“执行”的意思是“实施、实行”，是不是准确全面地反映了 executive 的意思呢？

一、执行功能的概念

Lezak(1983)指出 executive function(EF)是人们成功从事独立的、有目的的、自我负责的行为的能力，它包括目标形成、策划过程(具有抽象思维性质)、完成目标导向的计划和有效操作 4 种成分。它不同于其他各种如记忆、计算、语言、结构等认知功能。它要问的是一个人正在“怎样”做事情(如：他完成了这项工作，但他是如何完成的)，其他各种认知功能要问的一般是“什么”或“多少”(如：他对这件事了解多少，他能做什么)。若 EF 完整无缺，即使其他认知功能有相当大的损害，他依然能长久地维持独立的、积极的、负责任的行为；EF 受损后，不管他的其他认知功能保存得如何好或各种知识技能的得分如何高，他不再具有完好的自我照顾或正常的社交活动能力。EF 的损害涉及行为的所有方面，如情感平淡、易激惹、冲动、漫不经心、僵硬、注意力不集中和行为转移困难等很多行为问题来源于 EF 损害。从社会心理学的角度看，这些问题中最严重的是始动困难、动机缺乏，不能做成计划和活动次序等目标导向的行

为，接受和表达功能没有显著损害的患者易被误认为诈病、懒惰或宠坏、精神错乱或“获益性神经症”等。

1994 年 Sultzer 在“老年神经精神病学教材”中将 EF 区分为以下 4 种成分：① 动机：行为始动力。② 程序：模式识别、次序识别、交替选择。③ 反应控制：注意分割、抑制错误反应、认知速度、灵活性与计划性。④ 演绎推理：相似性和谚语理解等抽象思维、认知表达控制、反馈运用和预期能力。

Banich(1997)将 EF 障碍分为 5 类：① 心理惰性，如自发语言、自发行为减少，将意向付之行动有困难，出现环境依赖综合征。② 抽象思维能力减退，如能按照颜色(知觉)将动物卡片分类，不能按“驯服-野蛮”特性归类。③ 认知评估障碍。④ 处理新信息、应付新情境能力减退。⑤ 有目标导向的行为如次序排列、定势转移、策略修改、自我控制等障碍。

根据以上描述，EF 不仅仅是外在的“执行”，它更多的是指形成执行内容的动机、抽象推理、情势评估、灵活应对等内在思维活动。但是，约定俗成，作者仍然将 EF 译为“执行功能”。

二、额叶与 EF 的关系

EF 损伤最先是在额叶损伤患者中发现。医生们发现额叶损伤的患者智力、记忆、常识等与发病前相同，但是行为产生了变化，解决问题的能力低下，并且在做卡片分类等需要执行能力参与的神经心理测验时成绩低下，故把 EF 笼统地定位在额叶。但是，晚近研究认为，EF 不仅仅定位于额叶，它同其他大脑皮质、皮质下结构及小脑都有关联，也就是说，不能将 EF 与额叶功能等同起来。前额叶皮质(prefrontal cortex，PFC)不是一个单一的结构，它可以再分为 3 个主要区域，三者协同整合。

1. *外侧前额叶皮质*(dorsolateral prefrontal cortex，DLPFC)　起源于海马的旧皮质趋向的一部分，包括 Brodmann 分区的第 44、45、46 区的大部分，第 9 和 10 区的外侧，负责行为的认知部分，如计划、认知控制，是空间和逻辑推理过程等执行功能的结构基础。DLPFC 病灶的患者有严重的障碍，这些复杂认知任务包括需要发现规则、转移心理定势、解决多步骤问题、抵抗环境干扰、分配注意资源、维持和操纵非自动化反应、选择已加工的信息、积极提取记忆信息等等。这些认知加工的大部分依赖工作记忆，工作记忆指在短时间中维持和运用外部世界的相关信息，并使用这些内部表征详细计划目标导向的行为的能力。DLPFC 位于感觉联络区和运动系统的交界处，DLPFC 接受相关的

感觉信息、经由反馈通路过滤混杂的感觉刺激。DLPFC 与海马形成区的连接使当前的心理表征与以往的经验相比较，起辅助决策作用。而且，从其他边缘结构输出的情感和动机信息对于根据行为奖惩情况调整认知计划是必不可少的。另一方面，DLPFC 与运动系统和基底节的密切联系意味着 DLPFC 参与执行和控制反应。DIPFC 神经元的生理特征对于该区域在工作记忆和计划方面发挥作用也颇有意义。

2. 眶和腹内侧前额叶皮质　包括 Brodmann 分区的第 11、12、13、24、25 和 32 区的大部分，第 10 区的内侧，负责界定总体目标，根据上下文背景调整行为。患者则不能够评估现实情境和自己行为的未来结果。在人类的扑克牌游戏中，一组是点数较大、短期赢但长远可要输，另一组是点数较小、短期输但长远可要赢，结果，正常人选择后一组牌子，而腹内侧前额叶皮质的患者选择前一组牌子。DLPFC 和“眶和腹内侧前额叶皮质”是否存在相互作用？fMRI 研究显示奖励预期提高了参与认知加工的额叶的激活水平，但是，奖励预期的强度和认知加工的复杂性达到一定程度才能激活“眶和腹内侧前额叶皮质”，也就是说，存在一把“情绪阀门”过滤产生不良后果、干扰注意资源的情绪信息。

3. 背内侧前额叶皮质　包括 Brodmann 分区的第 9 区的内侧大部分，邻近前扣带回皮质，在前 SMA（辅助运动区）和 SMA 的后侧，负责行为反应的启动和不同可能反应冲突的处理。前额叶病灶患者可观察到的第 3 种损害是淡漠。它包括控制内生性行为（如监督和自我调节目前正在从事的行为的潜在冲突）出现更多的困难，这些行为与前扣带回皮质相关。这种功能障碍的一个解释是前扣带回皮质连接了前额叶皮质和纹状体，而且，前扣带回皮质涉及调控情感和行为的自主神经功能影响。

三、EF 分因子在大脑皮质的定位

早先 EF 被视为单一的认知结构。Baddeley 在 1986 年提出了 EF 的工作记忆模型，认为工作记忆由 3 部分组成：语音环路、视空间模板和中央控制器。后者代表执行功能或者是额叶功能。Baddeley 还认为 Morman 和 Shallice 的注意管理系统（SAS）也是中央控制器的模型。迄今已经有很多文章报道了不同群体样本（正常人或脑损伤患者）进行的一系列公认的执行功能测验，如威斯康星卡片分类测验（Wisconsin card sorting test，WCST）、汉诺塔测验（tower of Hanoi，TOH）、Stroop 色词测验等，用相关回归分析和因子分析，发

现各个测验的相关性很低(相关系数少于0.4)。Miyake和同事将EF划分为记忆刷新、定势转移和优势抑制3个因子,但是他们运用的测验中双重任务测验和这3个因子没有相关性。文献中还有一些其他的EF的分因子,如流畅性、概念形成、干扰控制、计划和组织、警惕性、估计等等。各个因子间并非独立,Miyake等人的研究表明,抑制、转换和刷新3种EF成分间主要以抑制功能为主。这3个分因子之间既存在着相互的联系又有相互独立的一面。Collette等沿用Miyake的结论用PET来研究3个成分的定位,结果发现右侧顶内沟、左侧顶上回、左外侧前额叶皮质均有激活,也证明了EF分因子有着相互的联系。

1. 优势抑制(inhibition)因子　将注意力集中在相关信息及处理过程上,抑制无关的信息和不适宜的优势反应。测试抑制最经典的神经心理测验是Stroop色词测验。维度变换的卡片分类、手部游戏、go-no-go任务等也可以用来测试抑制。Nelson认为不同的抑制任务激活不同的区域:解决刺激潜在冲突的时候,额下回激活,而解决刺激回答冲突时前扣带回激活。D' Esposito解决词语工作记忆干扰任务时候左额下回激活。这个结果和一些Stroop测验的皮质激活研究相类似。Fabienne CoUette的研究结果显示3个因子都激活的部位是右额下回(BA45)、右眶额回(BA11)、右中上额回(BA10),相关性分析显示相对转移和刷新来说这些部位同抑制关系更加密切。

2. 定势转移(set shifting)因子　是内源性注意控制机制,当两项任务竞争同一认知资源时,对两项任务相互转换的控制过程。常用的神经心理测验是双作业任务和连线测验(TMT)。研究显示,这个分因子和额叶、基底节有关。Smith等设计心理测验在排除了工作记忆成分干扰后发现前额叶背外侧皮质连同顶颞叶部分区域同转移有直接关系。Moll等运用口语连线测验,排除了该测验视空间和运动灵活性的干扰,发现前额叶背外侧皮质(BA6外侧、BA44和BA46区)辅助运动区和前扣带回显著激活。顶叶在转移当中的作用也渐渐被人们所重视,Rushworth等2001年研究发现视觉注意转移和视运动转移都激活顶叶,之后有许多人都报道了顶叶在转移中的作用,包括顶内沟、顶上回、顶下小叶等。Wager(2004)meta分析中描述了各式注意转移(地点转移、规则转移、物体转移、任务转移)的7个独立的区域——包括后部(顶叶枕叶)和前部(额叶背外侧皮质和前岛叶)这些区域在所有转移任务当中都激活。有人报道额叶损伤患者的转移能力并不受到太大影响而顶叶受损受到的影响大。这都说明顶叶在转移中扮演的角色可能比前额叶更重要。许多研究中都

发现前扣带回在转移中也起作用，而 Dreher 进行了双任务测验和任务转移测验，发现前者激活了前扣带回、后者激活左外前额叶和双顶内沟，从而提出扣带回是解决刺激反应关系中的冲突而外侧额叶主要处理转移过程。Collette 的研究发现 Miyake 定义的刷新定位在额极（BA10），上、中、下（BA6、BA9/46、BA44/45）回，眶额部（BA11）、顶内侧回和小脑，且以左侧的额极为主。

3. *流畅性（Fluency）因子*　要求受试者就某一语言或非语言范畴列举尽可能多的例子，检测命名能力、言语生成速度、短时和长时记忆，也反映受试者的语义组织和提取策略，后者反应执行功能。用得最多的就是类别流畅性作业（说动物的名称）、字母流畅性、语音流畅性。另外还有图形流畅性（如五点测验）等。许多人用 fMRI 研究了流畅性的定位，发现左额下回（BA44，45）、额中回（BA46，9）、前扣带回、岛叶、颞上回、小脑和它有关，并且以左侧为主。有文献报道语义流畅性以 BA45 区为主，而 BA44 区更多的是管理言语的组织过程，对于可能的皮质下结构如尾状核、豆状核和丘脑的激活可能是和发音清晰度、发音动作及语言处理有关，如果改良了言语流畅性测验，消除发音对于测验准确性的干扰，则皮质下激活消失。

总之，执行功能并非一个单一功能，它存在分因子，各个分因子之间有关联却又相对独立。它们在大脑中有各自独立的又有相同的定位区域，这个区域十分广泛，涉及前后大脑、基底节及小脑，并且相互联系，但以前额叶背外侧皮质和顶叶皮质为主。

（郭起浩）

第二章 神经心理测验总论

第一节 记忆门诊要掌握的神经心理测验

由于出生率减少、预期寿命延长，社会人口老龄化，老年人口不断增加，老年认知障碍就诊者不断增加。根据流行病学调查，60 岁以上老人中痴呆的患病率约 5%，轻度认知损害（MCI）的患病率约 15%，这是非常巨大的数字，高于脑卒中的患病率，所以，在全国的三甲医院，大部分开设了“记忆门诊”或“痴呆门诊”，有少数直接叫“神经心理门诊”，这些门诊不同于普通的神经内科门诊，就诊者绝大部分有认知减退的主诉或者是内科、神经科、老年科转介而来，在这些门诊，摆在医生案头的除了张明园教授主编的《老年期痴呆防治指南》和贾建平教授主编的《中国痴呆与认知障碍诊治指南》，还面临的问题就是：选择那些心理测验作为门诊基本工具？

如果说一般神经科医生只要掌握 MMSE、画钟测验等基本技能就足以应对临床需要，那么，作为记忆门诊显然是不够的，因为记忆门诊不仅要对痴呆及其分类作出判断，而且要对 MCI 与 VCI 及其亚型作出诊断，即对痴呆的危险人群及危险程度进行识别与区分。

由于每家医院记忆门诊的人员配置不一样（有的有专职的评定员，有的有兼职的评定员，有的没有评定员）、目的不一样（有的侧重分子生物学基础研究，有的侧重明确诊断，有的侧重照料指导），所以，采用的神经心理评估工具也不同，但是，有些工具应该是每个记忆门诊都必须具备的。让我们先看看欧美国家的记忆门诊平时采用的是哪些神经心理测验。

一、一些著名调查与研究采用的神经心理测验

（1）Weintraub 等 2005 年调查美国 29 个 AD 中心采用的神经心理测验的使用情况。最常用的是简明精神状态量表、动物流畅性测验、连线测验和波士顿命名测验（表 2－1－1）。

表 2－1－1 美国 29 个 AD 中心采用的神经心理测验

使用百分比	中文名称	英文名称
≥80%的中心使用	动物流畅性测验	Category Fluency (Animals)
	波士顿命名测验	Boston Naming Test
	简明精神状态量表	MMSE
	连线测验	Trail Making Tests (Part A or B)
≥60%～79%的中心使用	临床痴呆评估量表	CDR-Global Score
	字母流畅性测验	Letter Fluency Test
	韦氏智力测验-数字广度	Digit Span (WAIS-R)
	韦氏记忆测验	WMS (original, R or III)
	老年抑郁量表	Geriatric Depression Scale
	画钟测验	Clock Drawing Test
40%～59%的中心使用	CERAD 词表学习测验	CERAD Word List Learning
	Blessed 痴呆量表	Blessed Dementia Scale
	韦氏智力测验-搭积木	Block Designs (WAIS)
	数字符号转化测验	Digit Symbol
	神经精神量表	Neuropsychiatric Inventory
	韦氏记忆测验-视觉再生	Visual Reproduction (WMS)
	CERAD 全套神经心理测验	CERAD Neuropsychological Battery
<40%的中心使用	Blessed 常识-记忆-注意测验	Blessed Information-Memory Concentration
	加利福尼亚词语学习测验	California Verbal Learning Test
	Mattis 痴呆评定量表	Mattis Dementia Rating Scale
	总体恶化量表	Global Deterioration Scale
	新成人阅读测验	New Adult Reading Test
	阿尔茨海默病评估量表	Alzheimer's Disease Assessment Scale
	Lawton-Brody 日常生活能力量表	Lawton-Brody Activities of Daily Living
	Beck 抑郁清单	Beck Depression Inventory
	霍普金斯词语学习测验	Hopkins Verbal Learning Test
	Buschke 选择提醒测验	Buschke Selective Reminding Test
	Rey 听觉词语学习测验	Rey Auditory Verbal Learning Test
	认知功能电话问卷	Telephone Interview of Cognitive State
	Fuld 物品记忆测验	Fuld Object Memory Test
	修订 MMSE	Modified Mini Mental State Examination
	流调中心研究用抑郁量表	Center for Epidemiologic Studies-Depression Scale

(2) Maruta 等 2010 年调查了 34 个欧盟国家的会员，给予反馈的有 25 个国家，213 个神经心理测验被提到，其中 104 种测验只在某一国家使用而在别的国家不使用。所有国家都使用的测验有 4 种：MMSE、连线测验、言语流畅性测验和画钟测验。其他常用的包括临床痴呆评估量表(clinical dementia rating scale, CDR)、听觉词语学习测验(auditory verbal learning test, AVLT)、复杂图形测验(complex figure test, CFT)、符号数字转换测验(symbol digit modalities test, SDMT)、Stroop 色词测验(Stroop color and word test, SCWT)、波士顿命名测验(Boston naming test, BNT)、威斯康星卡片分类测验(Wisconsin card sorting test, WCST)、韦氏智力量表(Wechsler intelligence scales, WAIS)、韦氏记忆量表(Wechsler memory scales, WMS)、日常生活能力量表(activity of daily living scale, ADL)、神经精神量表(neuropsychiatric inventory, NPI)、Beck 抑郁清单(Beck depression inventory, BDI)和老年抑郁量表(geriatric dpression scale, GDS)等。

这说明，不同地区并没有统一的神经心理评估工具，但均追求简明扼要、适用性和针对性强的工具，基本上没有国家采用 Luria-Nebraska 神经心理测验(Luria-Nebraska neuropsychological battery, LNNB)、Halstead-Reitan 测验(Halstead-Reitan battery, HRB)这些著名的成套神经心理测验。

出于早期诊断、长期随访的目的，MCI 及其更早的“主观认知缺损(subjective cognitive impairment, SCI)”的评估势在必行，那么，MCI 的评估采用的是哪些神经心理测验呢？除了根据各种神经心理测验的品质，如信度和效度，我们还可以从国外权威的研究和指南中获得启发。

(3) 美国 AD 神经影像学研究(The Alzheimer's Disease Neuroimaging Initiative, ADNI)是 2004 年 10 月开始，为比较 AD 的各种临床前指标(包括系统的心理测验、1.5T MRI、FDG-PET、PiB-PET、血遗传学指标和脑脊液检查)的预测价值，开展主要针对 MCI 的 6 年随访研究，其设计严谨、检测手段先进、耗资巨大。入组 229 例正常老人、402 例 MCI、188 例轻度 AD，美国 57 个研究单位参加，迄今已经发表基于该研究数据的高质量论文 200 余篇，其研究模式被欧洲、大洋洲、日韩所仿效。

ADNI 采用的成套神经心理测验包括 MMSE、逻辑记忆-Ⅰ与Ⅱ、动物流畅性测验、数字广度、连线测验、SDMT、BNT、AVLT、画钟测验(自发与模仿)、阿尔茨海默病评估量表、GDS、NPI、CDR、ADL 等 14 种。

(4) 2006 年美国国立神经疾病和卒中研究所-加拿大卒中网络方案提出

血管性认知功能损害(vascular cognitive impairment, VCI)的 60 min 神经心理测验,包括:① 动物流畅性测验;② 受控口语词语测验;③ WAIS-Ⅲ的数字符号编码测验(digit symbol-coding subtest);④ 连线测验;⑤ 霍普金斯词语学习测验(或加利福尼亚词语学习测验);⑥ 简单和复杂反应时;⑦ BNT;⑧ Rey-Osterrich CFT;⑨ NPI;⑩ 流调中心研究用抑郁量表(CESD);⑪ IQCODE;⑫ MMSE。

(5) 2011 年出版的 AD 源性 MCI 的神经心理测验(Albert 等,2011)包括:记忆测验(选择提醒测验、加利福尼亚词语学习测验或 Rey 词语学习测验、逻辑记忆测验)、韦氏记忆测验之视觉再生分测验、连线测验、BNT、字母流畅性测验、动物流畅性测验、Rey-Osterrich CFT、数字广度测验。

与 ADNI 成套神经心理测验比较,这个协助 MCI 诊断的方案特点是:除了经典的语言情景记忆的评估,强调非语言情景记忆的评估(视觉再生与 CFT 回忆测验、ADNI 均没有);视觉空间的评估不是用画钟测验(CDT),而是采用难度更高的 CFT 模仿。

二、笔者总结的记忆门诊选择测验的 3 个等级

综合以上介绍,可以确定国内记忆门诊在选择测验方面可以区分为 3 个等级。

(1) 基本配置:MMSE、记忆测验(听觉词语学习测验或逻辑记忆测验)、动物流畅性测验、数字广度测验、画钟测验、连线测验、日常生活能力量表和老年抑郁量表等 8 种,是相应认知领域中最常用、最简单、最容易获得常模数据的测验,总耗时约30 min,反映总体认知、记忆、语言、注意、空间、执行等领域。

(2) 加强配置:有条件的有志于开展认知障碍研究的单位还应该加上临床痴呆评估量表、数字符号转化测验、波士顿命名测验、Rey-Osterrich 复杂图形测验、Stroop 色词测验和神经精神量表等 6 种,耗时增加约 40 min。不管是基本配置还是加强配置,一般需要一定程度的培训,有些测验需要购买测验版权。

(3) 探索性配置:当然,如果要走在学科前沿,这些常规的、标准化的评估工具还是不够的,应该增加探索性的认知检查,如社会认知与行为调整的检查、范畴特异性的检查等,根据每个研究者的特长与兴趣自行选择(表 2-1-2)。

表 2-1-2 记忆门诊配置神经心理测验的不同层次

领域	基本配置	加强配置	探索性配置(举例)
总体认知	简明精神状态量表	临床痴呆评估量表	病前智力
记忆	听觉词语学习测验或逻辑记忆测验	R-O复杂图形测验回忆	多重记忆系统
语言	动物流畅性测验	波士顿命名测验	范畴特异性
空间	画钟测验	R-O复杂图形测验模仿	虚拟空间
注意	数字广度测验	数字符号转化测验	注意选择与分割
执行	连线测验	Stroop色词测验	社会认知与行为调整
其他	日常生活能力量表老年抑郁量表	神经精神量表	照料者干预

基本配置已经能够满足MCI与VCI的诊断及其亚型区分的要求,为什么在基本配置的基础上再增加检测项目?这是因为单一测验检查未必能够全面准确地反映相应领域的功能状况,不同测验相互参照可以更恰当地反映相应的认知功能领域。当一个认知领域有多个分析指标时,阳性认知领域可以是多个指标同时下降、部分下降、任一指标下降、根据权重合成一个总分下降,等等,可以根据研究的侧重点确定操作性判断标准。

相比之下,目前国内使用比较多而又没有"上榜"的测验包括:汉语失语检查法、韦氏智力量表第一版之中文修订版(WAIS-RC)、韦氏记忆量表第一版之中文修订版(WMS-RC,不同于WMS-R)、威斯康星卡片分类测验(WCST)等,这些测验尽管经典、普及、知名度高,但它们有自己的适用范围,用于老年认知障碍门诊尤其是MCI的检查针对性不足、耗时过长、部分分测验的敏感性与特异性比较差,不在推荐之列。

第二节　应用神经心理测验判断认知损害

神经心理学测验经常用于神经科、精神科、老年科的临床实践和研究、干预措施制定及疗效评价等领域。当一位主诉"健忘"的老年人就诊时,医生首先应该想到的是记忆力的客观检查。包括韦氏成人智力量表修订版(WAIS-R)、临床记忆量表(CMS)、听觉词语学习测验(auditory verbal learning test,AVLT)及Rey-Osterreich复杂图形测验(Rey-Osterreich complex figure test)等,并借此判断受试者是否存在记忆损害、记忆受累的环节或成分及受累程

度；在分析检测结果时，受试者及其家属往往提出异议，认为医生的评价过于宽松或过于严格而不够准确。此时我们一般会告诉受试者或其家属检测结果仅供参考，就像血液白细胞计数对于感染的意义一样，确切的临床诊断还需根据患者的神经功能、影像学和实验室检查结果进行综合判断。因此，临床医生深刻理解各项神经心理学测验对认知损害的检测及分析方法，有助于更好地使用神经心理学测验方法。

常用神经心理学测验检测认知功能是否损害及其严重程度的方法有3种，即与常模(norm)比较、纵向随访比较及与估计的发病前认知水平比较。

一、与常模比较

常模是一种可供比较的普通形式。通常有如下方法。

1. 均数　运用某一受试者所测得的成绩(粗分或称原始分)与标准化样本的平均数相比较，以确定其成绩的高低。

2. 标准分　均数所说明的问题较为局限，只看均数，不注意分散情况，所得受试者的信息十分有限。如用标准分作常模，便可提供更多的信息。标准分能说明受试者的测验成绩在标准化样本的成绩分布图上居何位置。标准分(Z score)等于受试者成绩(X)与样本均数($\bar{\chi}$)之差(即 $X-\bar{\chi}$)除以样本成绩标准差(SD)，即 $Z=(X-\bar{\chi})/SD$。由此，不仅说明受试者的成绩与样本比较所在的位置，而且还能说明相差标准差的多少。韦氏成人智力量表(WAIS)采用的离差智商(deviation intelligence quotient，DIQ)的具体公式为，$IQ=100+15Z$(Z为标准分)。

3. T分　T分是标准分衍化出来的另一种常用常模。例如，明尼苏达多重性格调查表(Minnesota Multiphasic Personality Inventory，MMPI)便采用这种常模。它与离差智商的不同之处是所设的均数及标准差不同。

4. 百分位法(percentile rank，PR)　它的优点是不需要统计学的知识便可理解。习惯上将成绩差的排列在下，好的在上。计算出样本分数的各百分位范围。将受试者的成绩与常模相比较，如位于百分位数50(P50)时，说明受试者的成绩相当于标准化样本的第50位，也就是说，样本中有50%成绩在其之下(其中最好的至多和他一样)，另外50%的成绩优于该受试者。

5. 划界分(cut off score)　在筛选测验中常用该常模。如在教学上采用百分制时，以60分为及格分，此即划界分。在临床神经心理学测验中，将正常人与脑疾病患者的测验成绩相比较，设立划界分，用此分数划分有无脑损害。

如果某项测验对检测某种脑损害敏感。即说明设立的划界分十分有效。患者被划入假阴性的人数就很少甚至没有，正常人被划为假阳性的亦很少或没有。如果不敏感，则假阳性或假阴性的机会均会增加。

常模比较标准对于大多数心理学研究的目的来说是有用的，包括儿童与成人认知功能的描述、教育和职业规划及人格测评或评价。对一些可疑的行为若能利用常模比较标准与之比较，可直接对认知障碍进行评价。然而，当确诊或怀疑患者的认知功能(与成年人群中常态分布比较)减退时，除有发病前认知水平的相关书面文件(例如在校期间的成绩单)外，仅根据常模对其进行描述，就其本身而言对于评价受损程度的意义不是十分显著。获得的平均水平对于发病前就显得迟钝的患者，并不能说明所考核的认知水平显著下降；相反对发病前能力卓越的受试者而言，则可能代表存在认知障碍。

测量成人认知损害的第一步是建立患者发病前所有认知功能和能力的绩效水平，或者在无法获得直接信息时给予估计。对于那些有着大样本分层的常模，这一步简单易行。成人不能再命名物体、不能模仿一个简单的图案或对一侧身体产生忽视是明显的认知障碍。由于常模比较标准是反映正态分布的认知功能和能力的平均水平，因而只有个体化的比较标准才能对评价和鉴别认知障碍提供有意义的基线。对于单一患者，人群平均水平不是一个恰当的、必不可少的比较标准。从定义出发，对任何精心设计的、分数呈正态分布的神经心理学测验，都将有一半人能达到平均水平；各种结果相对于平均水平都存在差值。

针对认知改变的评价应该提倡个体化比较标准，无论其心理学特征或认知功能和能力在完全正常人群中是否呈正态分布都应作为一条法则。这条法则普遍适用于认知障碍和行为改变的测量。在测量的能力有各种各样或惯用的常模时(如手指敲击速度或听觉辨别精度)，常模比较标准对脑损伤的测量是适用的。即使所测量的这些能力随年龄发生改变、在绩效水平上存在性别差异而需要人群标准化，也同样适用。

二、纵向随访比较

个体化比较标准可以在仅依赖于对个体纵向随访进行前后比较的变化率研究中得到最好的例证。同一组测验应间隔一段时间进行，测验 3 次或更多次后进行测验评分时间序列上不同的比较。在儿童心理学研究中，变化率的

测验对于检测其发展率是不可或缺的。退行性疾病患者所表现的减分率信息有助于提高预测其病程发展的准确性。在以康复为目的的患者中，脑损伤后的认知功能改善率可能不仅对预测最终绩效水平有益，而且还可为实施康复的有效性提供信息。而且，变化率的研究有助于理解脑损伤对智力的长期影响。

对于在人群中普遍呈正态分布的能力，对受试者认知损害的认定取决于那些从历史资料（包括以前的测验评分）中获得的、可被假定为其发病前的特征性认知水平与现有测验表现（定量和定性特点）之间的比较。因此，许多临床神经心理学测验将能力和技能的个体内部比较考虑在内。

使用个体化比较标准可使认知障碍的检测成为简单而直接的操作过程：检查者需对可疑行为发生前与最近的情况，以及评估期间的差异进行比较。Hoofien 等对退伍军人应征入伍时的韦氏成人智力量表修订版评分和大脑受伤约 13 年后的评分进行比较，其结果为颅脑创伤患者存在认知障碍提供了明确的证据。

使用个体化比较标准的直接方法需要获得发病前的测验分数、学校成绩或其他相关的观察性资料。许多情况下，这些资料不存在或很难获得。因此，检查者常需从个体化比较标准中推断出间接的认知障碍评价方法。当前，在老年退行性疾病的早期诊断研究中，国际性的前瞻性随访调查资料不断增加，表明个体变化率的比较研究已日益受到重视，我国也应加强这方面的研究工作。

三、与估计的发病前认知水平比较

在间接测量方法中，检查者对患者目前的表现与原有认知水平的估计值进行比较，这些估计值可从许多来源获得。检查者的任务是寻找有意义的、正确的资料，从而对外伤前或发病前的认知水平进行估计，作为患者自身的比较标准。采用不同的方法推断出每例患者的比较标准，已经在各项临床应用中取得了不同程度的成功。显然，历史性及观察性资料是对患者发病前认知水平进行直接估计的信息来源。根据这一来源所得估计值的满意程度取决于对患者过去了解的多少，以及所知道或获得的患者的特点是否足以与其他人相区别。例如，检查者对某例颅脑创伤后认知损伤患者的了解仅限于他是一位受过 9 年教育的伐木工人，他的词汇水平和兴趣与其职业和受教育程度相符，则检查者仅能估计该患者的认知水平恰好处于平均水平，并以此作为比较标

准;而患者以前的智商比大部分人高,逻辑推理优秀或讲故事出色,或者脑损伤前即将被提升为主管等信息,检查者很可能难以获得.他们无法仅从患者的历史和观察中了解到该伐木工人以往的智力水平。

由此可见,仅根据以往和现场观测的数据对受试者发病前的认知水平进行估计,可能造成估计值偏高;而且,部分患者的主诉也可能高于实际水平。准确地对受试者发病前认知水平估计的需求显得越来越重要,特别是对老年人智力减退方面进行的估计。因此,神经心理学家设计了许多与众不同的评价方法以回应这一需求。

1. 智力测验评分　采用智力测验评分评价发病前的认知水平测验评分技术的基本特点是从分数本身评价受试者发病前的认知水平。长期以来,一直为临床所用的评价发病前认知水平的方法是,采用词汇评分作为唯一最佳的原始智力水平的指标(Yates,1954)。这种方法基于以下观察:许多认知功能减退的患者近期发生记忆、推理、数学运算能力及其他认知功能严重损害。然而很久以后仍能保留早年习得的、良好的语言技巧。并且,韦氏成人智力量表修订版所有的分测验中词汇评分与受教育程度的相关性最强。受教育程度同时也可作为发病前认知水平的一项良好指标。然而,词汇测验,如韦氏成人智力量表修订版需要口头定义。与可用一二个词语回答、只需再认或根据实际经验的语言功能测验效果相比。它对脑损伤的评价作用较差。而且,许多左侧大脑半球损伤致语言技巧受损的患者都表现为语言功能测验(>1 个)评分相对较低。其中以失语症患者的语言障碍最为明显,有些失语症患者完全不能使用语言符号。从技术上分析,部分左侧大脑半球损伤者并造成失语,但他们言语流畅性的下降足以使词汇评分无法提供良好的比较标准。

2. 韦氏成人智力量表　应用相同的原则划分"退纯比率"主要是比较词汇和其他用于检查注意力缺陷及视觉运动迟缓的测验中负载的言语部分的评分。基于大部分脑损伤患者仍然保留某些认知技巧的假设。结合个体的韦氏成人智力量表修订版与人口统计学变量(年龄、性别、种族、受教育程度、职业)完成一系列逐步回归方程,可以发现常识和词语是检测言语智商(verbal intelligence quotient,VIQ)和全量表智商(full-scale intelligence quotient,FSIQ)的最佳估计值,积木、拼图和填图分测验是操作智商(performance intelligence quotient,PIQ)的最佳估计值。

3. 单词阅读测验　用于评价发病前认知水平,尝试改善基于词汇的各种

方法。使之能应用于评价伴有各种痴呆性认知障碍患者的测验中。主要有以下几种。

(1) 国家成人阅读测验(national adult reading test,NART):其分数能够可靠地评价患者发病前的认知水平。NART需要患者口头朗读50个发音不规则的单词。这些单词的使用频率各不相同。当然,该项测验仅适用于像英语这种由许多单词构成、存在不规则发音的语言。本质上,这些单词阅读测验提供了一个词汇量的统计值。NART转换成智商(intelligence quotient,IQ):其分数的估计值与韦氏成人智力量表修订版(英国版)的全量表智商的相关系数为0.720~0.810。它与言语智商的相关性更高一些。与操作智商的相关性则相当低。Crawford等对179例11岁时曾参加过团体智能测验(Group Mental Ability Test)(推测是由纸笔方式完成)的77岁老年受试者施行NART和韦氏成人智力量表修订版(英国版)测试,结果显示NART估计值与智商分数的相关系数为0.730,后期的估计值与早期的测验评分在同一范围内。

(2) 北美成人阅读测验(the north american adult reading test,NAART):包括61个单词,其中35个来源于原始的NART。NAART分数与韦氏成人智力量表修订版言语智商的相关性良好($r=0.830$),与全量表智商的相关性亦较好($r=0.750$)。但存在大量未被解释的变量,与操作智商的相关性较差($r=0.400$),以至于不能作为有效的发病前认知水平的总体指标。

(3) 同音异义字的意义生成测验(homophone meaning generation test,HMGT):为一项对大脑前部组织损伤十分敏感的检测技术,已应用于患者认知功能损害程度的评价。此项任务需要认知定势的转移,向受试者提问一组即8个同音异义字的其他含义(例如pear-pair vs pare;sight-site vs cite)。Crawford和Warrington(2002)的观察发现,HMGT与NART这两项测验间的相关性具有统计学意义($r=0.605, P<0.05$)。由此推导出一个公式,即通过评价HMGT原始分数与NART估计出的发病前分数的差异,将很可能为"认知障碍严重程度"提供最佳的估计值。

(4) 大范围能力测验-阅读测验(reading test of the wide range ability test,WRAT-READ):系根据与NART测验同样的原则发展起来的阅读测验。它使用出现频率由多到少的单词(不是所有的WRAT-READ单词在发音学上都是不规则的)评价受试者的阅读水平。对WRAT-READ评价发病前认知水平的有效性方面的研究结果类似于NART及其变量的研究。

上述这些阅读测验与韦氏成人智力量表测验中智商分数的相关性，趋于与受试者的受教育程度直接相关。然而，当受试者的年龄范围横跨几个年龄段直至老年时，其以年龄作为影响因素的作用方能显现出来。尽管年龄的影响在大范围的受试样本(17～88岁)中十分显著($r=-0.178$)，但当受教育程度($r=0.510$)和社会地位($r=-0.355$)这些呈强相关性的影响因素被剔除后，年龄的影响作用即变得很小甚至趋近于零。

总之，对于认知功能完善的受试者，根据阅读水平预测言语智商程全量表智商分数，其测验结果十分准确。无论使用何种韦氏成人智力量表，NART/NAART或WRAT-READ评分与言语智商之间都呈相关性；虽与全量表智商的相关性略差，但仍能使大部分的变异得到解释；而与操作智商的相关性较差，以至于阅读测验评分不能对相关方面进行任何预测。此外，实际的智商分数越偏离100分，由NART或其变量中的一个所得的估计值的偏差就越大：在预测智商的应用范围扩大中存在截尾现象，个体能力水平的分布低于或高于平均水平均会导致估计值不可靠。进一步改善对发病前认知水平的估计产生了单词再认测验分数与人口统计学变量相结合的公式。

以上介绍了英语国家应用神经心理学测验对发病前认知水平进行评价的研究结果。鉴于我国由于方言复杂、个体受教育程度差异巨大，编制全国性的类似的阅读测验存在较大困难。但是，这项基础性研究工作无疑是有意义的，有助于我们更好地理解受试者的测验结果。

(郭起浩)

第三节　轻度认知损害的标准化神经心理测验

轻度认知损害(mild cognitive impairment，MCI)一词是1982年Reisberg等在编制认知功能障碍分级量表即总体衰退量表(global deterioration scale，GDS)时首次使用的，他们将认知功能和社会职业功能有轻度损害，但日常生活无明显影响的老年人归为MCI。此后，比较一致的看法是：MCI是正常衰老与痴呆之间的过渡状态。

针对老年人痴呆前状态的认知障碍，文献中曾经有很多术语，如年龄相关记忆损害(age associated memory impairment，AAMI)、年龄相关记忆减退

(age related memory decline, ARMD)、年龄相关认知减退(age related cognitive decline, ARCD)、良性老年遗忘(benign senescent forgetfulness, BSF)、非痴呆认知损害(cognitive impairment no dementia, CIND)、轻度认知障碍(mild cognitive disorder, MCD)、轻度神经认知障碍(mild neurocognitive disorder, MND)、可疑痴呆(questionable dementia, QD)、亚临床认知损害(subclinical cognitive impairment, SCI)等。10 余年来，基本统一为 MCI。

一、MCI 的诊断标准

Petersen 等于 1999 年首先提出的 MCI 临床诊断标准，包括"有记忆减退的主诉、有记忆减退的客观证据、总体认知功能未受影响、日常活动能力正常和非痴呆"五个方面。其作为遗忘型 MCI 的诊断标准，目前仍然得到广泛应用。随着研究的深入，人们发现 MCI 可以涉及众多认知域，而不仅仅有或一定有记忆损害。所以，Petersen 于 2004 年对 MCI 诊断标准做了修订并进一步提出 MCI 可以区分为四个亚型。同年 MCI 国际工作组提出了 MCI 广义诊断标准及诊断流程，诊断标准包括：① 认知功能障碍，但未达到痴呆的诊断标准(不符合 DSM-Ⅳ 和 ICD-10 的痴呆诊断标准)；② 认知功能衰退，患者和(或)知情人报告且客观检查证实存在认知损害，和(或)间隔一段时间检查发现有认知功能减退的证据；③ 基本日常生活能力保持正常，复杂的工具性能力可轻微受损。该标准不再强调将记忆损害作为 MCI 的诊断必备条件，并提出复杂工具性能力在 MCI 患者的变化。值得注意的是，它提到了随访的重要性。两年后，欧洲阿尔茨海默病协会(EADC)MCI 工作小组确立的 MCI 概念及诊断程序与上述标准相似。

客观的认知功能检查：神经心理测验证实存在客观认知功能损害，通常采用"比年龄和教育匹配正常人群常模低 1.5 个标准差(SD)，即小于等于均数 −1.5SD"。有个别研究(Busse, 2006)认为采用"≤1.0SD"获得的 MCI 有更高的预测价值。有的研究(如 ADNI-Ⅱ)认为：把 MCI 作为一个连续谱，早期 MCI，客观记忆检查得分在标准化测验常模的 1.0～1.49SD；晚期 MCI，客观记忆检查得分在标准化测验常模的 1.5SD 以上。这种区分的一个好处是把主观认知缺损(SCD)与 MCI 连接起来了，因为 SCD 的定义要求客观神经心理测验的得分在 1.0SD 以内。

有关 MCI 向痴呆转化的神经心理学与生物标志物研究已经汗牛充栋。

表 2-3-1 总结了各种 MCI 诊断手段的优缺点。

表 2-3-1　各种 MCI 检查方法的优点和缺点

类　别	指　　标	优　点	缺　点
认知测验	情景记忆，如词语延迟回忆、故事延迟回忆、联想学习；语义记忆，如语义流畅性、名人面孔识别；执行功能，如心理加工速度	易接受、易获得	临床前患者不够敏感
结构影像学	MRI 容积测量；颞叶内侧视觉评估量表；脑萎缩程度；弥散加权 MRI	易接受、较高敏感性	特异性偏低
功能影像学	SPECT 扣带回和左额叶区血流量、PET 颞顶叶区葡萄糖代谢、fMRI 功能网络分析	易接受、较高敏感性	特异性偏低
分子影像学	PiB-PET 等	敏感性和特异性高	费用高，设备依赖
电生理学检查	EEG 反映的 θ、α、β 活动、事件相关电位	易接受、易获得	敏感性和特异性偏低
脑脊液	Aβ 与 tau 蛋白检测	敏感性和特异性高	创伤性，接受差

二、新的 MCI 诊断标准

最近的研究认为 AD 诊断可以划分为三个阶段：第一阶段是“临床前 AD (preclinical AD)”，患者已经有生物学指标改变，是最早期的信号，在这个阶段，还没有临床诊断标准；第二阶段是“AD 型 MCI 或预期发展为 AD 的 MCI (MCI due to AD)”，患者的记忆和思维能力轻度改变，能够被观察到、被评估，但是日常生活和功能没有受损害；第三阶段是“AD 型痴呆(dementia due to AD)”，患者的记忆、思维和行为症状改变已经损害患者的日常生活和功能。

2011 年出版的美国国立衰老与阿尔茨海默病协会推荐的 MCI 诊断标准，将 MCI 诊断标准区分为核心临床标准(core clinical criteria)和临床研究用标准(clinical research criteria)。前者的定义与诊断标准已如本节描述，后者结合了生物学指标，仅用于发病机制和药物临床试验的研究中。由于生物学指标及其正常值不是每个单位都容易获得的，临床研究用标准还不能被推广普及。

临床研究用 MCI 标准将 MCI 试用性地区分为 3 种类型。

1. 很可能 MCI　受试者符合 MCI 核心临床标准，同时，分子生物学指标和神经损伤指标均呈阳性，该患者发展为 AD 有“最高的可能性”，因此，这部

分患者被确认为“很可能 AD 型 MCI”诊断。

2. 有可能 MCI　受试者符合 MCI 核心临床标准，反映 Aβ 沉积的指标阳性，而未检测到神经损伤，或者相反，神经损伤指标阳性而反映 Aβ 沉积的指标未被检测。由于生物学指标检测不全，随着时间推移发展为 AD 的可能性是中等的，因此这部分患者被称为“有可能 AD 型 MCI”。

3. 不发展为 AD 的 MCI　反映 Aβ 沉积和神经损伤的指标均为阴性，未来发展为 AD 的可能性最低，但是，这种 MCI 患者仍然有患 AD 的可能性，其病因值得进一步研究。

表 2-3-2　结合生物学指标的 MCI 诊断

诊断范畴	AD 病因指标	Aβ(PET 或 CSF)	神经损伤(tau, FDG, 结构 MRI)
MCI-核心临床诊断	没有信息	矛盾/不确定/未证实	矛盾/不确定/未证实
AD 型 MCI-有可能	中等	阳性	未证实
		未证实	阳性
AD 型 MCI-很可能	最高	阳性	阳性
MCI-不发展为 AD	最低	阴性	阴性

MCI 的症状不都是 AD 病变的结果，而有很多种病因，比如帕金森病(PD)、脑血管病、癫痫等。识别 MCI，并找到潜在的病因，对选择正确的治疗方法非常重要。针对不同病因的 MCI，其评估的认知领域的范围、神经心理测验的选择与受损的确认等有各自的特点。

三、AD 型 MCI 的情景记忆评估

虽然很多研究已经使用生物标志物(如脑脊液 Aβ 42 水平)来预测 MCI 患者的认知下降或进展为 AD 型痴呆的过程，但是其仍有很多限制。首先，确定这些生物标志物的截断值和诊断准确性尚在进行中(尚无国际共识)；其次，医院的条件与患者的接受度影响这些检查手段的普及；再次，即使已经完成生物标志物的检查，对如何合理应用这些指标评估认知下降和进展的认识也不足。故 MCI 的神经心理测验仍然是非常有用的检查方法。

AD 型 MCI 的神经心理测验有助于了解 AD 的病理过程，纵向认知评估结果不仅对于诊断的准确性很重要，而且有利于评估患者对治疗的反应。测验通常涉及 5 个认知领域：注意、记忆、空间、语言、执行功能。针对每个认知领域有经典测验推荐。其中情景记忆损害是最具 AD 预测价值的指标，是 AD 前驱期核心症状。文献中已经报道的、用来评估情景记忆的常用方法有：听

觉词语学习测验(auditory verbal learning test, AVLT)、逻辑记忆测验(logical memory, LM)、Rey-Osterrieth 复杂图形测验(complex figure test, CFT)、词语配对联想学习测验、韦氏记忆测验修订版(Wechsler memory scale revised, WMS - R,1987)、阿尔茨海默病联合登记组织(Consortium to Establish a Registry for Alzheimer Disease, CERAD)采用 CERAD 的 10 词语回忆分测验(CERAD - CWL,要求对 10 个词语进行长时延迟回忆)(Shankle, 2005)。美国 NIA - AA 2011 年推荐的情景记忆测验有 5 种：自由和线索选择性提醒测验、Rey 听觉词语学习测验(RAVLT)、加州言语学习测验(CVLT)、逻辑记忆测验(LMT)、非语言材料(如 Rey 复杂图形之延迟回忆)回忆。

针对情景记忆评估的操作性层面,我们有许多疑问：首先,相对于自由回忆,线索选择回忆是否为更佳选择;其次,听觉词语学习测验有即刻回忆、延迟回忆、线索回忆、选择回忆等步骤,哪个步骤更适合作为代表性指标;再次,是选择其中的一个记忆测验还是几个测验组合?

表 2 - 3 - 3 是笔者目前使用的 AVLT 版本(操作步骤见第四章第三节),10 多年来已经积累了 3 千多例评估资料,其中对部分遗忘型 MCI(aMCI)每年进行随访评估,以 aMCI 向 AD 的转化率作为主要判断标准、以 aMCI 的识别率、非 aMCI 向 AD 的转化率(即遗漏率)及向正常认知状态的转化率(即逆转率)作为辅助判断标准,分析不同测验、不同指标,作为 aMCI 识别的操作性诊断依据,获得以下结论：延迟回忆优于即刻回忆,结合结构影像学检查证实延迟回忆得分(而不是即刻回忆得分)与海马萎缩程度最相关;对于高龄老人,长延迟回忆与短延迟回忆具有等效性,而低龄老人长延迟回忆的效力优于短延迟回忆;AVLT 单一测验比组合测验(如 AVLT+复杂图形回忆)更有利于筛选患者与预测转化;短延迟回忆比长延迟回忆更有利于海马型与非海马型痴呆的鉴别诊断。

长延迟回忆由于过度敏感而可能出现假阳性,故线索回忆与元记忆部分可以辅助长延迟自由回忆指标。线索回忆有五种：① 1 对 1 类别再认,如回忆“锣鼓”一词时,问刚才学习的词语中乐器是哪一种;② 1 对多类别再认,如表 2 - 3 - 3 中的 N6;③ 多选再认,如 回忆“锣鼓”一词时,问刚才学习的词语是“二胡、钢琴、锣鼓”中的哪一个;④ 首字母再认,如果是回忆英文单词,可以问刚才学习的词语中“B”开头的词语是哪一个;⑤ 是/否再认,如表 2 - 3 - 3 中的 N7。目前认为,就其线索强度来说,类别再认比较弱而是/否再认最强,多选再认比较适中。

表 2-3-3　听觉词语学习测验(AVLT)的 8 次回忆

	即刻回忆			延迟回忆			线索回忆			元记忆
词语	N1	N2	N3		N4		N5	N6 (类别线索)	N7(是/否判断)	N8 - JOC
1				间隔非语言测验5分钟		间隔非语言测验20分钟		花朵类		“请你自己估计一下，刚才针对24个词语的判断，你说对了多少个词语?”
2										
3										
4										
5								职业类		
6										
7										
8										
9								服饰类		
10										
11										
12										

元记忆有四种：① 任务难度判断(ease of learning，EOL)；② 学习程度的判断(judgement of learning，JOL)；③ 提取自信度判断(judgement of confidence，JOC)；④ 知晓感判断(feeling of knowing，FOK)。表 2-3-3 中的元记忆检查是否有助于 SCD 或 MCI 的诊断，正在分析中。

综上所述，笔者采用 AVLT 判断 AD 所致 aMCI 的情景记忆受损时，最合适的客观指标是 AVLT 长延迟回忆得分下降，同时参考即刻记忆指标、线索回忆指标与元记忆的得分情况。

四、PD 所致 MCI 的认知评估

PD 所致 MCI(PD-MCI)在非痴呆患者中很常见，占 19%～38%。PD-MCI 由临床、认知和功能标准共同定义，其临床特征具有异质性，可有很多认知域的损害。总的来说，非遗忘型、一个认知域受损是最常见的 PD-MCI 亚型。诊断 PD-MCI，必须首先临床确诊为 PD，PD-MCI 以由潜在的 PD 进展引起的认知功能下降为特点。认知下降由患者或者知情者报告，或者由临床医生发现。认知缺陷必须有客观测试证据。最后，认知缺陷不会显著影响功能独立性。根据 Litvan 等 2012 年发表的 PD-MCI 指南（Mov Disord，2012，27(3)：349-356)，关于神经心理测验部分介绍如下。

PD-MCI 的诊断标准根据神经心理测试的综合结果分为两种可操作的

水平。水平1和水平2都用来检测PD-MCI,但在评估方法、诊断确定程度和临床特征上不同。

(一) 水平1: 筛查

水平1基于简略的认知评估,诊断确定性低于水平2。水平1要求在总体认知功能评估(表2-3-4)中有认知损害,或者在有限的神经心理测试中表现出认知损害(如1个认知域只有1个测试,或者测试少于5个认知域)。相关的认知域包括注意与工作记忆、执行功能、语言、记忆、视空间能力(表2-3-5)。当使用的是有限的神经心理测试时,诊断水平1的PD-MCI必须至少有2个测试呈现出认知损害。

表2-3-4 评估总体认知能力和病前智力的神经心理量表

评 估	神经心理测试	测试时间(min)
总体认知	蒙特利尔认知评估	10
	PD认知评定量表	15
	帕金森病认知量表	15
	Mattis痴呆评定量表	20～30
可评估的病前智力	国家成人阅读测验	5
	韦氏成人阅读测验	5

表2-3-5 各认知域测验推荐及测试时间

认 知 域	神 经 心 理 测 试	测试时间(min)
注意与工作记忆	数字字母排序测验(来源WAIS)	5
	连线测验	5～10
	数字广度测试或数字编码	5
	Stroop色词测验	5～10
执行功能	Wisconsin卡片分类测验或尼尔森改编版	15
	剑桥神经心理自动化成套测验	10～15
	言语流畅性测验	5
语言	相似性测验(来源WAIS)	10～15
	波士顿命名测验(或PD简略版)	
	分级命名测验	5～15
记忆	词表学习测验,比如Rey听觉言语学习测验、加利福尼亚词语学习测验、霍普金斯言语学习测试和选择性提醒测验	10～20
	散文回忆测验,比如逻辑记忆测验(来源WMS)或Rivermead行为记忆试验短语回忆分测验	10～15
	简略视空间记忆测验修订版	10～15

（续表）

认知域	神经心理测试	测试时间(min)
视空间能力	本顿线方向判断测验	5～10
	霍普视觉组织测验	10
	画钟测验(如 Royall 的 CLOX)	

注：WAIS，韦氏成人智力量表；WMS，韦氏记忆量表。

(二) 水平 2：全面评估

水平 2 建议使用正式的、全面的神经心理测试，5 个认知域中每个认知域必须至少测试 2 个。要求同一个认知领域有两个测验受损或两个不同认知领域各有一个测验受损。认知损害可以表现为多种形式：与年龄、教育程度、性别匹配的正常老人相比，得分低于 1～2 个 SD，或连续认知测验表现出显著下降，或与病前智力相比显著下降。PD-MCI 分型：PD-MCI 单认知域——1 个认知域的 2 个测试异常，其他认知域未受损；PD-MCI 多认知域——2 个或者更多认知域各自至少 1 个测试异常。

对于 PD 相关认知损害的分子和影像生物标志物已有研究，但大多数研究缺乏可靠数据。因此，目前并不推荐将生物标志物纳入 PD-MCI 的诊断标准，但如果将来生物标志物的可靠性大大提高的话，可能会重新修订诊断标准。

复旦大学附属华山医院神经内科王坚教授领导的运动障碍专病组于 2013～2015 年在 PD 患者中验证了上述神经心理测验，发现需要肢体操作的测验(如连线测验、数字符号转换测验、画钟测验)不适合 PD 患者。通过大量预试验，笔者与王坚一起开发以下测验组合作为标准化的中国人 PD-MCI 的成套测验(见表 2-3-6)。

表 2-3-6　中国人 PD-MCI 的成套测验(王-郭编制)

领域	测验	测验时间(min)
总体	简易智能状态检查(MMSE)	5
	蒙特利尔认知评估基础量表(MoCA-B)	15
记忆	听觉词语学习测验(AVLT)	8
	数字序背测验(DOT)	2
注意	数字顺背与数字倒背测验(DS)	2
	符号数字转化测验口头版(SDMT-O)	2
语言	波士顿命名测验(BNT)	5
	动物流畅性测验(AFT)	1

（续表）

领　域	测　　　　验	测验时间(min)
空　间	线方向判断测验(JLO)(15 项)	5
	物品识别与判断测验(VOSP)剪影分测验	5
执　行	Stroop 色词测验(24 个词)	5
	Go/no－go 测验	1
		约 56

五、血管性 MCI

长期以来，脑血管病被认为是认知损害的一个重要原因，诊断血管性 MCI 首先要确定存在轻度认知损害，再确定血管疾病是引起认知缺陷的突出的但非绝对唯一的病因。

脑卒中(stroke)是一种突然起病的脑血液循环障碍性疾病，又叫脑血管意外。它是指因各种诱发因素引起脑内动脉狭窄、闭塞或破裂，而造成急性脑血液循环障碍，临床上表现为一过性或永久性脑功能障碍的症状和体征。脑卒中分为缺血性脑卒中和出血性脑卒中。高达 64％的脑卒中患者存在某种程度的认知功能损害，有 1/3 会发展为明显的痴呆。相应地，尸检病理研究提示 34％的痴呆病例存在显著的病理性血管改变。而且使个体罹患脑血管疾病的因素同样也是导致认知功能损害的危险因素。

由血管因素导致或与之伴随的认知功能损害被称为血管性认知功能损害(VCI)。VCI 可单独发生或与 AD 伴发。而且，脑血管病理与 AD 病理间似乎存在很强的相互作用，同时有两种病理改变的患者的认知功能损害比只有一种病理改变者更明显。因为大部分血管性危险因素是可以干预的，所以 VCI 和伴有血管因素的 AD 是可以预防或可延迟、缓解其发展的。早期有关脑血管疾病相关认知功能损害的诊断与治疗只聚焦于痴呆即血管性痴呆(VaD)这种严重的认知功能损害。最近，那些表现为非痴呆 VCI(VCI－ND)的患者得到重视，并被认为是临床试验的最佳对象，因为他们的认知功能损害尚处于轻微阶段。

VCI 的大体分类见表 2－3－7。

最近发表的专家共识(Sachdev P 等。Alzheimer Dis Assoc Disord，2014，28(3)：206－218)把血管性 MCI 称为轻度血管性认知障碍(VCD)，VCD 的异质性使得发现可靠的非影像学生物标志物极具挑战性。一些被建

表 2-3-7 VCI 的分类

严重度分类	病因学分类	疾 病
VCI-R(VCI 危险人群)*	心血管	冠状动脉旁路移植术、心脏骤停、急性心肌梗死
	脑血管	一过性缺血发作(TIA)、慢性偏头痛
	全身性	高血压或严重低血压、糖尿病、低血氧-低灌注性(如血容量不足、失血性休克)、肥胖症
	精神性	血管性抑郁
VCI-ND(非痴呆血管性认知损害)	大血管缺血性	皮质性脑梗死、多发性脑梗死、关键部位梗死等
	小血管缺血性	Bingswanger 病、伴有皮质下梗死和白质脑病的常染色体显性遗传脑动脉病(CADASIL)、腔隙性脑梗死等
VCI-D(血管性痴呆)	出血性	脑出血、蛛网膜下腔出血、脑淀粉样血管病、慢性硬膜下血肿等
	脑静脉血管病性	脑静脉窦血栓形成、脑动静脉畸形等
	血管炎	变态反应性血管炎、感染性血管炎(如结核性、梅毒性血管炎)
VCI-M(混合性痴呆)**	退行性痴呆+血管性因素	AD+血管性因素、FTLD+血管性因素、LBD+血管性因素、PDD+血管性因素、其他退行性痴呆(MSA、PSP、CBD、海马硬化)+血管性因素
	血管性痴呆+退行性因素	血管性痴呆+退行性因素

注：* VCI-R 指目前没有认知损害或只有轻微认知损害(达不到 VCI-ND 的程度)，但是病情发展可以出现 VCI。** 狭义的混合性痴呆指 AD+血管性因素。

议的 CVD 的 CSF 生物标志物有：白蛋白指数作为血脑屏障受损的标志物；硫酸酯作为脱髓鞘的标志物；神经丝作为轴索变性的标志物；基质金属蛋白酶作为血管疾病的标志物。因为没有一个是 VCD 所特有的，故它们并不被推荐用于诊断。所以，神经心理测验是目前诊断 VCD 的主要手段。

与诊断其他原因认知损害不同的是，诊断血管性认知障碍需要评估 7 个认知域，即注意力和处理速度、额叶执行功能、学习和记忆、语言、视结构—知觉、实践—直觉—身体图式和社会认知(表 2-3-8)。

表 2-3-8 在血管性认知障碍中评估的认知域

1. 注意力和处理速度：持久注意、注意分配、选择性注意、信息处理速度
2. 额叶执行功能：计划、决策、工作记忆、反馈/纠错回应、新环境、习惯抑制、心理灵活性、判断
3. 学习和记忆：即刻记忆、近事记忆(自由回忆、线索回忆)、识别记忆
4. 语言：命名、表达、语法和句法、感觉性语言
5. 视结构—知觉：结构、视知觉和推理
6. 实践—直觉—身体图式：实践、直觉、右/左定向、计算、身体图式、面容识别
7. 社会认知：情绪性和社会性线索的识别、恰当的社会抑制、心理理论、移情

大量的文献都得出这样的结论，即VCD的显著障碍是在处理速度和额叶执行功能上，表现为信息处理减慢、工作记忆变差、定势转移能力下降。在轻度VCD中，信息处理减慢尤其突出，这要求必须使用定时测试。在VCD中，情景记忆常常受损，但是损害模式可能与AD者不同，信息提取比保持更容易受影响，故与词表的自由回忆相比，患者的再认任务表现比较好。

六、总结

比较AD-MCI、PD-MCI与血管性MCI的认知功能检查，我们可以发现：① 认知评估通常有两个层次：简易筛查与全面正式。前者更多的是临床医生的初步筛查，后者是研究所用的系统检测；前者是全科医生必须掌握的，后者则是专科医生必须掌握的。我们必须避免把简易筛查得分当作诊断MCI的“金标准”，也不能强求没有条件的医院做全套测验检查。② MCI的分型有助于更好地明确痴呆的前驱期，在AD型MCI首先是区分遗忘型与非遗忘型，在PD型MCI是区分单认知域型与多认知域型，在血管性MCI则建议区分执行型与非执行型。简易筛查测验无法实现这种区分，故有条件的单位应该进行包括5～7个认知域的全套测验。③ 认知领域与受损标准的确定依据病因不同而有差别，AD的生物标志物比较明确而有效，而PD-MCI与VCD的生物标志物的效力比较低，所以，后者对测验的全面性要求更高。④ 在欧美国家，每个认知领域最佳的测验工具与指标的选择是基本相同的，不同的研究小组之间具有可比性。国内在采用这些成套测验方案时需要进一步研究，应在采用相同测验材料的情况下，采集不同地区的测验信度、效度与地方性常模，从而更精准地诊断患者。

（郭起浩　赵倩华）

第三章

认知筛查量表

在痴呆的大样本流行病学调查和基层医院初步判断是否存在认知障碍的方面，筛查测验(screening test)由于成本低、耗时少(通常<15 min)、基本上不需要培训(非评估专业人员操作完成)，以及对于MCI和痴呆的识别与随访跟踪有一定的敏感性和特异性而得到广泛的应用，常用的痴呆筛查测验除了MMSE，见于文献的认知简短筛查方法还有Blessed定向-记忆-注意测验(Blessed orientation-memory-concentration, BOMC)、简短精神状态问卷(short portable mental status questionnaire, SPMSQ)、认知能力筛查量表(cognitive capacity screening examination, CCSE)、长谷川痴呆量表(Hastgawa dementia scale, HDS)、Mini-Cog(三词回忆+画钟测验)、七分钟痴呆筛查测验(7 minute screen, 7MS，包括言语流畅性、定向、图片回忆、画钟测验)、Rowland通用痴呆评估量表(Rowland universal dementia assessment scale, RUDAS)和记忆损害筛查(memory impairment screening, MIS)。常用的MCI筛查测验有Dem tect (Kalbe, 2004)、AB认知筛查测验(AB cognitive screening test, ABCS, Molloy, 2005)、蒙特利尔认知评估量表(Montreal cognitive assessment scale, MoCA, Nasreddine, 2005)、记忆变化测验(memory alteration test, M@T, Rami, 2007)和快速认知筛查测验(quick cognitive screening test, QCST, Guo, 2010)。这些方法的比较见表3-0-1，列举的是3个常用的认知筛查方法。

1. Dem tect Dem tect是Kalbe等研制的快速有效的MCI测试工具，选取了及时延时记忆、数字转换、语言流畅、操作记忆等子测试。项目包括10个单词2次学习后即刻回忆和延迟回忆；列举"超市物品"的流畅性；数字顺背和倒背；数字转换(number transcoding)，后者如"six hundred and eighty one"

表 3-0-1　常见认知筛查量表比较

项　目	低费用	耗时少	易操作	非专业	高敏感	高特异	协诊断	高一致	利随访	利鉴别
1. MMSE < 25	+	7	+	+				+		
2. Mini-Cog	+	5	+	+				+		
3. Clock Draw	+	2	+	+						
4. FAQ > 8	+	5	+	+				+		
5. IQCODE	+	5	+	+				+		
6. MCAS		20						?		
7. TICS	+	3	+	+				+		
8. DQ	+	5	+	+				+		
9. CamCog		20						+	+	
10. DMT	+	10	+	+	?	?		+		
11. MIS-T	+	4	+	+				+		
12. CWLoriginal		10			+	+		+		
13. CWLcomputerized	+	10	+	+	+	+	+	+	+	+

说明：? =may possibly be satisfied but could not find published data on the stated criterion.
(1) Mini-Mental State Exam (MMSE); (2) Mini-Cog; (3) Clock Drawing test (CDT); (4) Functional Activities Questionnaire (FAQ); (5) Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE); (6) Minnesota Cognitive Acuity Screen (MCAS); (7) Telephone Interview for Cognitive Screening(TICS); (8) Dementia Questionnaire(DQ); (9) cognitive subtest of the CAMDEX battery (CAMCOG); (10) Double Memory test (DMT); (11) Memory Impairment Screen for Telephone (MIS-T); (12) delayed free recall subtest of the CERAD 10-Word List (CWLoriginal).

应该转换为 681，而不是 600801。13～18 分代表认知正常，9～12 分为 MCI，0～8 分为痴呆，耗时 8～10 min。敏感性等同 CDR，适用于早期或轻度认知受损。对 MCI 患者的识别明显优于 MMSE，且不受受试者教育程度和年龄的影响。

2. AB 认知筛查(AB cognitive screen test, ABCS)　包括定向、重复单词、延迟回忆、画钟、语言流畅等 5 个认知子测试，总分 135 分。能敏感地区分正常人群、MCI、痴呆。Molloy 等研究结果表明在使用标准 MMSE 测试 MCI 和正常人时平均得分差别不明显，而使用 ABCS 时得分差异有统计学意义，较少受受试者教育程度和年龄的影响。

3. 记忆改变测试(memory alteration test, M@T)　是针对遗忘型 MCI 的筛查工具，包括语言片断和语义记忆测试，5 min 完成，快速简易。Rami 等对 400 名西班牙老年人群的前瞻性研究，结果表明 M@T 能有效地识别 MCI 和早期痴呆。

第一节　简明精神状态量表(MMSE)

简明精神状态量表(Mini-Mental state examination,MMSE)是 Folstein 等编制的用于评估认知功能的简易工具,当时是用来测验教育年限大于 8 年老年人的认知部分而非用以筛查痴呆患者的工具。然而,经过实践摸索,MMSE 逐渐用于筛查痴呆患者、判断认知损害的严重度并跟踪记录病情变化情况。由于 MMSE 容易操作、耗时少(5～10 min),它自 1975 年问世以来在国内外得到推广普及,目前已经有 100 多种语言版本。中英文 3 种版本项目见表 3-1-1、表 3-1-2、表 3-1-3。MMSE 满分为 30 分。

表 3-1-1　MMSE 中文版——张明园修订版

项　目	问　　题
定向	1. 今年的年份? ______年　2. 现在是什么季节? 季节______ 3. 现在是几月? ______月　4. 今天是几号? ______日 5. 今天是星期几?　6. 现在我们在哪个市(省)? 7. 你家住在什么区(县)?　8. 住在什么街道? 9. 我们现在是第几层楼?　10. 这儿是什么地方?
登记(词语即刻记忆)	11. 现在我要说三样东西的名称,在我讲完之后,请你重复说一遍,请你记住这三样东西,因为等一下要再问你的:"皮球、国旗、树木。"最多重复 5 次。以第一次回答记分。(1) 皮球____国旗____树木____(2) 皮球____国旗____树木____(3) 皮球____国旗____树木____(4) 皮球____国旗____树木____(5) 皮球____国旗____树木____
心算	12. 假如你有 100 元钱,花掉 7 元,还剩下多少?(在受试者回答后,不管对错)问,再花掉 7 元,还剩下多少? 如此一直算下去,直到减去 5 次为止。不要重复受试者的回答。93________86________79________72________65________ (注意:当患者忘记减去 7 后的数字,不能给予"93 再减去 7"这样的提示,若前一个答案错了,但据此而得出的下一个答案都是对的,只记一次错误。)
词语回忆	13. 刚才我请你记住的三样东西是什么? 皮球____国旗____树木____
语言能力	14. 请问这是什么?　手表________请问这是什么? 笔________ 15. 请照着这卡片所写的去做。 16. 请你说一句完整的、有意义的句子。记下句子________________ 17. 现在我要说一句话,请清楚地重复一遍:"四十四只石狮子。" 18. (访问员说下面一段话,并给受试者一张空白纸,不要重复说明,也不要示范):请用右手拿这张纸,再用双手把纸对折,然后将纸放在你的腿上。
结构模仿	19. 请你按样画图。不要解释图形。

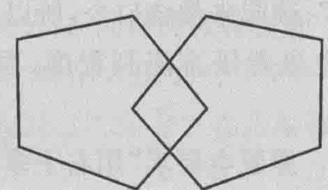
表 3-1-2 MMSE 香港版

项 目	问 题
定向	现在是什么年份？ 现在是什么季节？ 现在是什么月份？ 今天是什么日子？ 今天是星期几？ 你现在在什么国家？ 这儿是什么省份？ 这儿是什么区？ 你在什么医院？/你在什么地方？ 你在几楼？
词语即刻记忆	现在我要说出三种东西的名称。说完以后，请你重复一次，并记住它们。因为几分钟后，我会叫你再说一遍。[苹果]，[报纸]，[火车]。（以第一次讲的计分，一个一分，然后重复对象，直到全部三种都记住。）
心算	请你用 100 减 7，然后再减 7，一直减到我叫你停为止。（减 5 次后便停，93，86，79，72，65）
词语回忆	我刚才叫你记住的三种东西是什么？
语言能力	这是什么？（铅笔）（手表） 请你重复我说的句子 [姨丈买鱼肠]。 桌子上有一张纸。请你用右手拿起它，再用双手把纸折成一半，然后把它放在桌面上。 请你读出纸上的字，然后跟着去做。 请你说出任何一个完整的句子。例如：[我是一个人]，[今天天气很好]。
结构模仿	这里有一幅图画，请你照着画出来。

表 3-1-3 英文版 MMSE

Maximum Score	*Score*	*Questions*
		Orientation
5		"What is the year? Season? Date? Day? Month?"
5		"Where are we now? State? County? Town/city? Hospital? Floor?"
		Registration
3		The examiner names three unrelated objects clearly and slowly, e. g. "apple", "table", "penny" , then the instructor asks the patient to name all three of them. The patient's response is used for scoring. One point for each correct. The examiner repeats them until patient learns all of them, if possible. Count trails and record.
		Attention and Calculation
5		"I would like you to count backward from 100 by sevens." (93, 86, 79, 72, 65, ...) Alternative: "Spell WORLD backwards." (D-L-R-O-W)
		Recall
3		"Earlier I told you the names of three things. Can you tell me what those were?"
		Language Tests
2		Show the patient two simple objects, such as a wristwatch and a pencil, and ask the patient to name them.

（续表）

Maximum Score	*Score*	*Questions*
1		"Repeat the phrase: 'No ifs, ands, or buts.'"
3		"Take the paper in your right hand, fold it in half, and put it on the floor." (The examiner gives the patient a piece of blank paper.)
		Read and obey the following
1		"Please read this and do what it says." (Written instruction is "Close your eyes.")
1		"Make up and write a sentence about anything." (This sentence must contain a noun and a verb.)
1		"Please copy this picture." (The examiner gives the patient a blank piece of paper and asks him/her to draw the symbol below. All 10 angles must be present and two must intersect.) 
30		*TOTAL*

一、中英文两种不同语言版本的 MMSE 比较

MMSE 的中文译本很多，项目内容在不同中译本中略有差异。我们的分析依据首先是以 MMSE 原文为基础，然后结合张明园教授审定的目前最常用的中文版进行说明。

1. 定向　按照英文版本的要求，地点定向应该从最大单位（国）到最小单位（第几层楼），中文版从“省”开始，难度比“国”应该稍高；加拿大版本也是从“省”开始。日期和星期差一天即为错误。

2. 登记　也称即刻记忆、最初记忆或一级记忆，要求受试者复述并记住 3 个性质不同的物件，告知时需连续给出，应语音清晰、缓慢、大约一秒钟一个。重复学习最多 5 次。不要解释词语，如受试者问“shumu 是人道树的树木还是数目字的数目？”，可以回答“都行”，而不是告知“是人道树的树木”，后者为线索记忆。英文版 3 个单词是“苹果、便士、桌子”，应该比中文版的“皮球、国旗、树木”来得容易，另一方面，由于“便士”存在文化背景的影响，各国在翻译时往往不是对应翻译这 3 个词语。

3. 注意和计算　中文版本要求患者从100连续减7，每错一次扣一分。原版中有一个替换项目，倒说单词“WORLD”的字母，有的翻译为倒背“瑞雪兆丰年”，如倒背错为“年丰雪兆瑞”则为3分，以此类推。有的要求倒背数字(4-2-7-3-1)。由于老人教育程度的影响，往往不理解“倒背”的意义，故实际上运用不多。在“100连续减7”与“WORLD”倒背任务的选择中，一般的英文版本会选择后者，如ADNI研究方案采用的就是后者。然而，这2种检测方法不是等价的，后者的难度大于前者，一般受试者会先将该单词顺背，逐个拼写字母，再倒过来说。如果受试者对字母表不熟悉很可能得0分。

4. 语言复述　是检查语言复述能力，要求患者复述一中等难度的成语，Folstein原文为“no, ifs ands or buts”是一句成语，其意义是“不要为自己的失误找任何借口”，读起来像绕口令，所以，目前大部分中文版本翻译为一句绕口令，不仅检查患者语言流利程度，而且要求患者口齿清楚，难度比原版高。

5. 三步指令　原版本要求“用右手拿这张纸”，但是，有部分版本强调非利手，一般是左手，中文版“用右手拿这张纸，再用双手把纸对折，将纸放在大腿上。”仍然是“右手”。另外，细节上要求：口令没有间断，不理会打岔；用20 cm×30 cm白纸；非惯用手即不是拿筷子的手来拿纸；如果受试者说“没有听清楚”，可以说“只要做你听到的部分”；主试者不要做递纸和要纸的动作；如果折纸超过一次，对折部分就给0分。

6. 阅读理解　准备一白纸用粗体大字写“闭上您的眼睛”，请患者先朗读一遍，然后要求患者按纸写命令去做。患者能闭上双眼给1分。

7. 书写　Folstein原文要求给患者纸和笔，请患者在纸上主动随意写一个句子。考虑到中国老人中有一部分文盲不懂握笔，改口述句子代替患者自发书写。句子应有主语和谓语，必须有意义，能被人理解。香港版有举例：请你说出任何一个完整的句子。例如：“我是一个人”“今天天气很好”。

8. 临摹　要求患者临摹一交叉的2个五角形。两图形必须交叉，必须有10个角，交叉后的图需成四边形。但角不锐和边不直可忽略不计。

二、评分

1. 分析指标与正常值　MMSE分析指标为总分。英文版的最佳划界分

从21分到28分都有，比较常用的是25/26分，也就是＜26/30分。中文版MMSE通常依据不同教育程度制定划界分。张明园(1990)调查年龄在55～80岁城市社区人群，制定的划界分是：文盲组≤17分、小学组≤20分、中学或以上组≤24分，低于划界分为认知功能受损。张振馨(1999)通过大样本流行病学调查将划界分定为文盲组≤19分、小学组≤22分、中学或以上组≤26分。彭丹涛(2005)制定的非文盲人群中，年龄小于50岁者，划界分定为≤28分，大于80岁时划界分定为≤25分。在张明园的5年随访中，正常衰老的MMSE减少约0.25分/年，病理衰老约4分/年。

2. 中英文版本得分比较　根据前面的项目分析，有些项目是张明园中文版比英文版难，有些项目相反，修订后难度下降了。那么，这个中文修订版本在年龄、教育程度匹配的中美国家老人中具有可比性吗？表3-1-4与表3-1-5都是正常社区中老年人的得分，在同样的教育程度下，大致上，60～69岁、70～79岁、80～89岁，每增加10年MMSE总分下降1分左右，近乎匀速下降，如受教育“9～12年组”分别是28、27、26分。在张振馨的调查资料中(采用的是张明园修订版)，60～69岁、70～79岁、80～84岁也是近乎匀速下降，如“初中及以上组”，MMSE总分也是28、27、26分，说明MMSE的难度在2个国家中是接近的，因为在年龄与教育程度匹配的中美2组老人中MMSE得分非常接近。但是，张氏调查中，85～90岁组的得分有一个明显的“跌落”，这是不是一个普遍现象值得进一步调查研究，甚至可以将年龄范围扩大到100岁进行调查。

表3-1-4　北京城乡55岁或以上居民MMSE总分的分布情况

年龄(岁)	文盲组		小学组		初中及以上组	
	例数	均数(SD)	例数	均数(SD)	例数	均数(SD)
55～59	169	23.0(3.1)	394	26.4(3.2)	574	28.4(2.2)
60～64	388	23.1(3.8)	388	26.2(3.4)	600	28.4(2.1)
65～69	404	22.5(4.4)	354	26.3(3.6)	379	28.1(2.3)
70～74	331	21.8(4.8)	258	25.7(4.8)	234	27.6(4.2)
75～79	238	20.8(5.5)	130	25.5(3.4)	130	27.6(2.7)
80～84	140	19.3(5.9)	47	24.3(4.7)	50	26.4(4.7)
85～90	66	17.7(6.3)	32	21.7(6.7)	28	20.9(6.7)

资料来源：张振馨，洪霞，李辉，等. 北京城乡55岁或以上居民简易智能状态检查测试结果的分布特征. 中华神经科杂志，1999，32(3)：149～153.

表 3-1-5 美国不同年龄与教育水平的 MMSE 总分的分布情况

年龄(岁)	0～4 年组		5～8 年组		9～12 年组		大学及以上组	
	例数	均数(SD)	例数	均数(SD)	例数	均数(SD)	例数	均数(SD)
55～59	49	22(2.7)	208	26(2.9)	525	28(2.2)	231	29(1.5)
60～64	88	23(1.9)	310	26(2.3)	626	28(1.7)	270	29(1.3)
65～69	126	22(1.9)	633	26(1.7)	814	28(1.4)	358	29(1.0)
70～74	139	22(1.7)	533	26(1.8)	550	27(1.6)	255	28(1.6)
75～79	112	21(2.0)	437	25(2.1)	315	27(1.5)	181	28(1.6)
80～84	105	20(2.2)	241	25(1.9)	163	25(2.3)	96	27(0.9)
85～90	61	19(2.9)	134	23(3.3)	99	26(2.0)	52	27(1.3)

数据来源：美国 5 个地区(New Haven, Baltimore, St. Louis, Durham, Los Angeles)1980～1984 年的入户流行病学调查。

三、评价

1. MMSE 的影响因素　研究最多的是文化背景、性别、年龄、受教育程度和种族等。其中，教育程度对 MMSE 总分的影响最明显。因为文盲老人对于阅读、书写、计算等项目的影响是很明显的，有调查认为文盲的健康受试者的表现等同于受教育年限在 1～4 年的轻度痴呆患者的表现，所以在文盲人群中，MMSE 总分的诊断价值可能有限。在 Jacqmin-Gadda 等人的一项前瞻性研究发现年龄大及教育程度低的老人 MMSE 得分下降更明显。

虽然年龄对于 MMSE 总分的影响不如教育程度明显但也不可低估。注意力的检查可能在年龄相关的认知衰退更具特别意义。虽然有的学者提出 MMSE 中注意力的检查在不同的时间没有可靠的一致性。

性别对 MMSE 总分的影响不如教育程度明显，但是也有部分研究结果表明存在性别差异。例如加拿大研究者发现在平均教育水平为 10 年的老年人的 MMSE 总分，女性得分高于男性。而 Rosselli(2000)对教育年限大于 4 年的哥伦比亚老年人的研究发现男女之间并无显著的差异，但是教育程度低的(0～3 年)在男女之间是有显著差异的。Jones 和 Gallo (2002)报道在 MMSE 的部分项目是有性别差异的，如女性在连续减 7 表现会较差，而男性在倒背单词上困难一些。在国内，大部分研究则显示性别对 MMSE 总分没有明显影响。

其实，影响因素也是相互的，并非完全独立。在使用 MMSE 的时候，要注

意到选择的人群和使用的版本。在神经心理测量学中，通常要对分数进行调整以消除偏倚。Blesa (2001)等发现了经过年龄和教育程度的调整可以提高MMSE总分的特异度却不降低其敏感度。Mungas (1996)揭示了校准MMSE总分可以使得其在不同的文化、种族及教育程度人群中有较稳定的特异度和敏感度。还有另外一个办法补偿文化偏倚就是修正或者删除导致偏倚的项目。

2. 信效度　MMSE信度良好，联合检查的组内相关系数为0.99，相隔48～72 h重测，组内相关系数可达0.91。

MMSE的分析指标为总分，反映的是总体认知功能，它与各种其他测验有中-高度的相关性，比如，MMSE总分与WAIS的相关系数是0.78；与Mattis-DRS总分的相关系数是0.87；与ADAS认知分测验总分的相关系数是0.90；与画钟测验得分的相关性在0.82～0.85左右，后4者是综合性的认知检测工具。尽管Folstein(1975)在介绍MMSE的项目时，分别冠以“定向、登记、注意与计算、回忆、语言”等范畴，这只是小分段(subsection)，不是反映相应的认知领域(cognitive domain)或因子分(factor)。必须指出的是，不能将MMSE的单个项目得分或分段得分视作相应的认知领域表现，它与相应的神经心理测验的相关性非常低，如MMSE中两物命名不等同于命名能力，它与Boston命名测验得分并没有显著相关性；模仿画交叉五边形不等同于空间结构能力，它与Rey-Osterriche复杂图形测验的结构模仿得分也没有显著相关性。但是，最近有研究认为，定向、注意(连续减7)和记忆(3词回忆)一定程度上反映了相应的认知领域。总之，MMSE的项目分或分段分的分布不能作为该患者的详细的认知剖面图。当然，这并不意味MMSE在鉴别诊断方面一无是处，Brandt(1988)发现亨廷顿病患者注意项目比AD患者差而记忆项目比AD患者好，同样，Jefferson(2002)发现血管性痴呆与帕金森病在运动与结构、工作记忆(倒背WORLD与执行3步指令)项目比AD患者差，而定向与回忆项目比AD患者好。

MMSE用于随访是敏感的，每年下降2～5分，不同的研究者报告不一致，因为AD的认知功能下降不是匀速的，不同的认知领域下降速度不一样，个体之间也有差异。MMSE不适合重度患者的随访。针对重度痴呆患者建议采用DRS。MMSE用于预测痴呆也有一定的价值。针对早期患者，项目分析表明MMSE的3词回忆、100连续减7、模仿画图及时间定向对痴呆识别比其他项目更为敏感。MMSE总分与影像学脑萎缩程度、SPECT反映的脑灌注缺损

及事件相关电位的潜伏期延长有显著相关性。但是,深入研究认知损害不能单纯采用MMSE,须采用多个更特异的测验工具组合使用。临床医生不能仅依据低于MMSE总分的划界分作出痴呆诊断,就像不能单纯根据白细胞总数来认定是否感染一样,认知评定只能作为痴呆诊断的辅助工具,临床诊断必须结合患者的病史、日常活动能力变化、非认知行为症状及脑影像学、电生理学、血液和脑脊液生化学检查结果,根据相应诊断标准做出,最后确诊还有赖于随访、脑脊液检查和病理检查。

有些研究者通过对MMSE的项目增删试图改进诊断效度,较为常用的是认知能力筛查量表(cognitive abilities screening instrument,CASI),又称为3MS,包括定向、注意、心算、远时记忆、新近记忆、结构模仿、语言、范畴流畅性、概念判断等9个因子,共20题,费时15~20 min,总分100分,得分可换算成MMSE的分数,与MMSE相比,CASI并没有改善痴呆识别的敏感性。

3. 优缺点　MMSE在识别MCI方面是有局限性的。Mitchell(2009)总结5篇MMSE识别MCI的论文,阳性预测值(PPV)只有37.0%,也就是有超过60%的假阳性,阴性预测值(NPV)达到83.2%,所以,它被有些研究机构推荐作为诊断标准的一部分。事实上,MMSE不仅对MCI的识别意义有限,对轻度痴呆的识别也有不同意见,比如,Welsh(1991)报道MMSE识别轻度AD的敏感性是73%;Tombaugh(1992)报道MMSE识别早期痴呆的敏感性是21%~54%。

总而言之,MMSE有许多优点,比如,它耗时短,操作简便,易携带、易推广;在痴呆的起病和进展阶段均可使用;它的低分或下降速度可以作为痴呆预后的预测因素。MMSE作为痴呆诊断的辅助工具,敏感性高,尤其是在评估中、重度认知损害时假阴性率极低。所以,MMSE在社区大样本调查、药物临床试验及临床医生对可疑病例作初步检查等方面得到广泛的应用。

但是,MMSE的缺点亦不容忽视,表现在其项目内容易受到受试者教育程度的影响,对文化程度较高的老人有“天花板效应”,即可能出现假阴性,容易忽视轻度认知损害,而对低教育和受方言影响者则有可能出现假阳性;强调语言功能,非言语项目偏少;对右半球功能失调和额叶功能障碍不够敏感;没有时间限制;对皮质下功能紊乱不及皮质性功能紊乱敏感;用于不同病因所致痴呆的鉴别诊断的价值有限,因为PDD、VaD、FTD主要的缺损领域是执行功能,而MMSE恰好缺乏对额叶执行功能的评估项目。

最后要说明的是,使用MMSE是免费的,但是,MMSE的修订版,也就是

MMSE-Ⅱ，是需要支付版权费用的。从对 MMSE-Ⅱ的测验介绍看，新版本在项目组成上并没有根本变化，变的只是 3 个词语不一样等细节方面，所以，我们在临床与研究中可以继续使用 MMSE。

第二节 智能筛查测验(CASI)

智能筛查测验(cognitive abilities screening instrument，CASI)是一种简易认知功能检测工具，是美国南加州大学李眉(EL Teng)等编制的一套筛查痴呆的认知检查量表。本量表共 20 个题目，总分 100 分。能在 15～20 min 内对注意、心算、定向、旧记忆、新记忆、言语流畅性、语言能力、构图能力、概念判断作出定量评价。由于根据不同文化地理背景而修订某些题目，CASI 已形成一个系列，并以不同版本代号加以区别。其中 CASI C-2.1(表 3-2-1)为中文版，是专门根据我国文化背景编制，且适用于受教育水平偏低或未受过正式教育的受试者。CASI 已在美国、日本、中国大陆与台湾等试用。使用说明见表 3-2-2。

表 3-2-1 智能筛检测验(C-2.1 版)

序号	项 目	CASI 评分	MMSE 评分
1	你今年几岁？______	2 1 0	1
2	一年有几个月？______	2 0	
3	过年是几月几号？(12/30；1/1)	2 1 0	
4	一个钟头有几分钟？或：一年有多少天？(365；360 天)	2 0	
5	太阳是从哪个方向下山的？(可提供：东，南，西，北)	2 0	
6	月饼是什么节吃的？(中秋节)	2 0	
7	下面我要讲 3 个名词，你注意听好，记住这些词语。 在我讲完之后，你就照讲一遍[任选 1 组]： 1. 帽子______黄色______小孩______ 2. 鞋子______白色______邻居______ 3. 袜子______蓝色______朋友______	初次正确数 3 2 1 0	3
8	今年是哪一年？或：今年是什么年？	4 2 1 0	1
9	几月？(农___)______月	2 1 0	1
10	几号？(农___)______日	3 2 1 0	1
11	今天是星期几？______	1 0	1
12	现在是上午、中午、下午，或是晚上？______	1 0	1

（续表）

序号	项　　目	CASI 评分	MMSE 评分
13	想想看哪些动物有 4 条腿？告诉我愈多愈好。（30 秒）	0～10	
14	刚才我请你记的 3 个名词是什么？ 1）自动答出 2）“是一样穿的或是戴的”“是一种颜色”“是一个人” 3）“帽子，鞋子，袜子” “白色，蓝色，黄色” “朋友，小孩，邻居”	 1.5　1　0.5　0 1.5　1　0.5　0 1.5　1　0.5　0	 3
15	这里是商店店面、医院还是家里？________	1　0	1
16	这里是什么区（镇，乡，村）？________	2　0	2
17	这里属于哪一个市（县）？________	2　0	2
18	现在我要讲几个数目字，然后请你把它们倒念出来。譬如说，我说 1-2-，你就说 2-1-。开始： a. 1-2-3（如不对则教 3-2-1）[正确=1，错误=0] b. 6-8-2（如不对也不教）[正确=2，错误=0] c. 3-5-2-9（如 a=b=0，则跳过此项）	 1　0 2　0 2　0	
19	100 块钱用掉 3 块钱，还剩多少钱？　(97)再用掉 3 块，还剩多少钱？　(94)(再重复 3 次)91，88，85。	5　4　3　2 1　0	5
20	橘子和香蕉相同的地方是它们都是水果。（停 2 秒） ______和______有什么相同？______和______都是什么？ a. 鱼-虾　b. 桌-椅　c. 鞋-袜 d. 手-脚　e. 哭-笑　f. 吃饭-睡觉	只记 1，2，4 三题 6　5　4　3 2　1　0	
21	1）如果你邻居的房子失火了，你会怎么办？（每类一分） 2）如果你把借来的伞弄丢了，你会怎么办？（每类一分） 3）如果你在路上看到别人遗失的身份证，你会怎么办？	2　1　0 2　1　0 2　1　0	
22	请你仔细听我要讲什么，等我讲完，你就一字不差地照讲一遍。（停 2 秒） 他想要回去（2.5 秒） 下面请你照讲：（5 秒） 这个黄杯子（1　0）比红饭碗（1　0）还要重（1　0）	 2　1　0 3　2　1　0	 1
23	我想请你做一件事（出示卡片）请闭上眼睛 [会照着做=3；提醒了会做=2；会说不会做=1；读做都不会=0]	 1　0	 1
24	模范绘图（动笔时开始计时）○（15 秒）　◇（30 秒） （60 秒）交叉五边形	10～0	1
25	我想看看你写的字，请你写“人，父，母，子，女”	1	1
26	执行口头指令：请你用左（右）手来拿这张纸，把它对折一次，然后交还给我	3　2　1　0	3
27	先前我要你记住的 3 个名词是什么？ 1）自动答出 2）“是一样穿的或是戴的”“是一种颜色”“是一个人” 3）“帽子，鞋子，袜子” “白色，蓝色，黄色” “朋友，小孩，邻居”	 1.5　1　0.5　0 1.5　1　0.5　0 1.5　1　0.5　0	

（续表）

序号	项 目	CASI评分	MMSE评分
28	命名：额头___下巴___肩膀___手掌或掌心___大拇指___汤匙(调羹)______硬币(铜板，钱币)______牙刷______钥匙______梳子______	5 4 3 2 1 0	2 汤匙 硬币
29	请你记住这5样东西。(等5秒；盖住；再问：) 我刚才给你看的是哪5样东西？(每样0.6分) 匙______币______刷______钥______梳______ [可以告诉正确答案，但不要告诉还有回忆] 结束时间______测验耗时？分钟：______ 分数效度 (1) 有效；(2) 重听；(3) 视觉不良；(4) 动作不良；(5) 方言不通；(6) 神志不清；(7) 身(心)不适；(8) 其他原因	3.0 2.4 1.8 1.2 0.6 0 TIME	

注：1. 你好吗？(1) 好；(2) 不好；(3) 不好不坏；(4) 不会回答。
2. 如果回答不好，为什么不好呢？(1) 身体不适；(2) 情绪不良；(3) 其他原因。
3. 下面我想问你一些问题。有的很容易；有的比较难，很多人都不会。不会没有关系，你会的就告诉我，不会的就说“不知道”，好吗？

表3-2-2 智能筛检测验(C-2.1版)使用说明书

序号	项 目	不计CASI分	CASI评分	MMS评分
1	你今年几岁？______[差0～2岁：2分；差3岁：1分；差>3岁 0分]		2 1 0	1
2	一年有几个月？[正确2分；错误0分]		2 0	
3	过年是几月几号？(12/30；1/1)[正确2分；月或日正确1分；错误0分]		2 1 0	
4	一个钟头有几分钟？或：一年有多少天？(365；360天)[正确2分；错误0分]		2 0	
5	太阳是从哪个方向下山的？(可提供：东，南，西，北)[正确2分；错误0分]		2 0	
6	月饼是什么节吃的？(中秋节)[正确2分；错误0分]		2 0	
7	下面我要讲3个名词，你注意听好，记住这些词语。 在我讲完之后，你就照讲一遍：(3组任选一组) a. 帽子______黄色______小孩______ b. 鞋子______白色______邻居______ c. 袜子______蓝色______朋友______ 如第1次不能回答出，可重复并教最多3次。(此时不可再提醒受试者要记住)	最后正确数 3 2 1 0	初次正确数 3 2 1 0	3
8	今年是哪一年？或：今年是什么年？[正确4分；差1年2分；差2～5年1分；差≥6年0分]		4 2 1 0	

（续表）

序号	项　　目	不计 CASI 分	CASI 评分	MMS 评分
9	几月？（农____）________月[差≤5天，2分；差6天～1个月，1分；差≥2个月0分]		2　1　0	
10	几号？（农____）________日[正确3分；差1～2天2分；差3～5天1分；差≥6天0分]		3　2　1　0	
11	今天是星期几？________		1　0	
12	现在是上午、中午、下午，或是晚上？________ 11AM前：上午 11AM～12NOON：上午，中午 12NOON～2PM：中午，下午 2PM～黄昏：下午　黄昏后：晚上		1　0	
13	想想看哪些动物有4条腿？告诉我愈多愈好。（30秒） [记录全部回答，超过10个亦予记录，前15秒与后15秒回答用/隔开]	总分 错误数	0～10	
14	刚才我请你记的3个名词是什么？ 自动答出＝1.5；（3秒后）"是一样穿的或是戴的"＝1；（2秒后）"帽子，鞋子，袜子"＝0.5；仍不正确或不知道＝0 自动答出＝1.5；"是一种颜色"＝1.0；"白色，蓝色，黄色"＝0.5；仍不正确或不知道＝0 自动答出＝1.5；"是一个人"＝1.0；"朋友，小孩，邻居"＝0.5；仍不正确或不知道＝0 如未得满分（1.5，1.5，1.5）， 此时要再重述3个词一遍。	自动分 提示分 选择分	1.5　1　0.5　0 1.5　1　0.5　0 1.5　1　0.5　0	
15	这里是商店店面、医院还是家里？[若不在这3个地方，请在医院项代以正确地点]________		1　0	1
16	这里是什么区（镇，乡，村）？________		2　0	2
17	这里属于哪一个市（县）？________		2　0	2
18	现在我要讲几个数目字，然后请你把它们倒念出来。譬如说，我说1-2-，你就说2-1-。记住哦！ 要把我说的数目字倒念出来。 （每秒一字，可提醒"倒念"） a. 1-2-3（如不对则教3-2-1）[正确＝1，错误＝0] b. 6-8-2（如不对也不教）[正确＝2，错误＝0] c. 3-5-2-9（如a＝b＝0，则跳过此项）		1　0 2　0 2　0	
19	100块钱用掉3块钱，还剩多少钱？（97）再用掉3块，还剩多少钱？（94）（再重复3次）91，88，85。 注意：第1次回答错误时，给予正确答案；以后出错不予纠正；连续2次错误即终止。		5　4　3 2　1　0	5

（续表）

序号	项　目	不计CASI分	CASI评分	MMS评分
20	橘子和香蕉相同的地方是它们都是水果。（停2秒） ________和________有什么相同？________和________都是什么？ [6题都做，但是CASI只记1)、2)、4)题] 1) 鱼-虾 [海味，海鲜，水产=2分；食品，水中之物=1分；答错或不知道=0分] 2) 桌-椅 [家具=2；4条腿，木制，吃饭用=1；答错或不知道=0] 3) 鞋-袜 [穿在脚上的=2；穿的=1；答错或不知道=0] 4) 手-脚 [身体的一部分，四肢=2；五指(趾)，骨，肌肉=1；答错或不知道=0] 5) 哭-笑 [感情或情绪的表达=2；其他正确答案(如表情)=1；答错或不知道=0] 6) 吃饭-睡觉 [必需的身体功能=2；其他正确答案(如习惯)=1；答错或不知道=0] 第1和第4子题回答错误可以提供正确答案。	12分	只记1),2),4)三题 6 5 4 3 2 1 0	
21	1) 如果你邻居的房子失火了，你会怎么办？（每类一分，记录所有回答）[5项中任意2项] a. 报警，通知消防队，打“119” b. 通告或救屋内的人 c. 帮忙救火 d. 通知其他邻居 e. 保全自家人财 2) 如果你把借来的伞弄丢了，你会怎么办？（每类一分） 告知原主，道歉　(1　0) 赔偿　(1　0) [如果回答“我会去找”，不给分，但可以追问“如果找不到呢？”] 3) 如果你在路上看到别人遗失的身份证，你会怎么办？[交给警察、派出所；寄回=2；或其他适当有效作为；不去碰；其他次等行为=1；不适当之行为=0]		2 1 0 2 1 0 2 1 0	
22	请你仔细听我要讲什么，等我讲完，你就一字不差地照讲一遍。（停2秒） 他想要回去(2.5秒) 全对=2；错漏一二字=1；错漏更多字=0 下面请你照讲：(5秒) 这个黄杯子(1　0)比红饭碗(1　0)还要重(1　0)		2 1 0 3 2 1 0	1 第2句相当“44狮”
23	我想请你做一件事(出示卡片)请闭上眼睛 会照着做=3 提醒了会做=2 会说不会做=1 读做都不会=0	3 2 1 0	1 0	1

（续表）

序号	项　　目	不计 CASI 分	CASI 评分	MMS 评分
24	模范绘图(动笔时开始计时) ○ (15 秒) 大致封闭之图形=2;长短直径之比>2∶1=1;更差=0 ◇(30 秒) 四边并大致菱形=2;长短边之比>2∶1;其他封闭形=1;更差=0 (60 秒) 　　　　　　　每个五边形 大致正五边形　　4　　4 长短边之比>2∶1　　3　　3 其他封闭图形　　2　　2 不封闭,≥2 边　　1　　1 <2 边　　0　　0 　　　　　　　交叉 是四个角　　2 非四个角　　1 没有交叉　　0	2 1 0 2 1 0	DRAW 10～0	1
25	我想看看你写的字,请你写“人,父,母,子,女”(60 秒)[全部正确时,CASI=1 分]	5[每字 1 分]	1	1 [说一句话]
26	遵循三段口示 请你用左(右)手来拿这张纸,把它对折一次,然后交还给我 1) 口令没有间断,不理会打叉 2) 用 20 cm×30 cm 白纸 3) 非惯用手即不是拿筷子的手来拿纸 4) 如果受试者说“没有听清楚”,可以说“只要做你听到的部分” 5) 主试者不要做递纸和要纸的动作 6) 如果折纸超过 1 次,对折部分就给 0 分		3 2 1 0	3
27	先前我要你记住的 3 个名词是什么? (每项分数: 0,0.5,1,1.5) 1) 穿戴: 帽子,鞋子,袜子 2) 颜色: 白色,蓝色,黄色 3) 人: 朋友,小孩,邻居 [回答错误不用再告知]	自动分 提示分 选择分	0～4.5	

（续表）

序号	项　　目	不计 CASI 分	CASI 评分	MMS 评分
28	[脸上]这[部分]叫什么？（每部分2秒） 额头______下巴______肩膀______ 手掌（掌心）______大拇指______	部位 5 分	CASI 得分为 10 个名称的总分除以 2。<0 时记 0 分。 5　4　3　2 1　0	
	下面我要请你看几样东西，这是什么？[次序不管] 汤匙（调羹）______硬币（铜板，钱币）______ 牙刷______钥匙______梳子______	物品 5 分		
	如果受试者说不出，可以告诉他正确名称，要求受试者重复一遍。	重复（5—0）		
29	请你记住这 5 样东西。（等 5 秒；盖住；再问：） 我刚才给你看的是哪 5 样东西？ （每样 0.6 分） 匙______币______刷______钥______梳______		3.0　2.4　1.8 1.2　0.6　0	
	结束时间______：______ 测验耗时？分钟：______		CASI TIME	
	分数效度 有效 重听 视觉不良 动作不良 方言不通 神志不清 身（心）不适 其他原因		VLDT 1 2 3 4 5 6 7 8	

CASI 是由 MMSE 发展而来的，其项目难度与 MMSE 接近，其优于 MMSE 的地方是 CASI 有 9 个因子分，根据不同因子的缺损可以有助于 AD 的严重度判断和不同类型痴呆的鉴别诊断。

完成 CASI 即可同时得到 MMSE 分。但是，MMSE 与 CASI 中计算的 MMSE 评分还是有 4 个项目不同：① CASI 采用“100 连续减 3”×5 次，而 MMSE 系“100 连续减 7”×5 次；② CASI的地点定向是施测者地址，而 MMSE 是受试者家庭地址；③ CASI 是写 5 个字，MMSE 是说一句话；④ CASI命名调羹和硬币，MMSE 则是钢笔和手表。

Lin 等调查台湾地区 2 096 例 65 岁以上老人，制定的划界分是：教育程度 0 年、1～5 年和≥6 年的 3 组老人，CASI 总分的界定分分别是 50 分、68 分和 80 分。在作者的检测资料中，各组受试者的总体教育程度比较高（平均 12 年），正常对照组的 CASI 平均分约为 90 分，MCI 组平均约为 80 分，轻度 AD

组平均约为70分，中度AD组平均约为60分。

第三节　蒙特利尔认知评估量表(MoCA)

一、MoCA的意义

轻度认知损害(MCI)被认为是介于正常衰老和痴呆之间的过渡状态，对于痴呆的早期诊断或干预具有重要的意义。MCI的快速筛查建立在对其神经心理特征、诊断和分类的基础上。Petersen等于1999年首先提出的MCI诊断标准，该标准要求确定MCI必须有记忆主诉，最好有一个知情者确认；神经心理测验证实存在客观记忆损害，比年龄和教育匹配正常人群低1.5个标准差(SD)；总体认知功能基本完好；日常生活活动基本保存；没有痴呆。作者采用该标准诊断MCI，比较了不同记忆测验对MCI的识别率和转化率的差异，从而提出相对比较好的客体记忆检测方法。2004年MCI国际工作组和2006年发表的欧洲阿尔茨海默病协会(EADC)MCI工作组提出的MCI广义诊断标准包括：有认知功能障碍，但未达到痴呆的诊断标准；患者和(或)知情人报告且客观检查证实存在认知损害，以及(或)间隔一段时间检查发现有认知功能减退的证据；基本生活能力保持正常、复杂的工具性生活能力可有轻微损伤。

根据受损的认知领域，MCI可以有以下分类方法。

(1) 4个亚型：2个遗忘型和2个非遗忘型。遗忘型即：① 单纯记忆受损；② 多认知领域受损(记忆加1个或多个其他认知领域，如语言、执行、视觉空间领域)。非遗忘型：③ 单一非记忆领域受损；④ 多个非记忆领域受损。

(2) 4个亚型：① 单纯执行功能受损；② 包括执行功能多个领域受损；③ 单一的非执行功能受损；④ 不包括执行功能的多个认知领域受损(也就是用执行功能取代记忆功能作为分类主要坐标)。

(3) 3个亚型：① 遗忘型；② 非遗忘型；③ 记忆与非记忆认知功能同时受损。

(4) 3个亚型：① 执行功能受损；② 非执行功能受损；③ 执行与非执行功能同时受损。

(5) 2个亚型：① 单一认知领域受损；② 多个认知领域受损。

(6) 2个亚型：① 遗忘型；② 非遗忘型。

(7) 2 个亚型：① 执行功能受损；② 非执行功能受损。

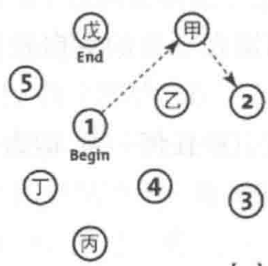
(8) 2 个亚型：① 1～2 个领域受损；② >3 个领域受损。

目前大部分专家认为前两种 4 个亚型的分类方法比较符合研究需要。

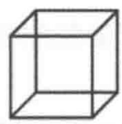
蒙特利尔认知评估(Montreal cognitive assessment，MoCA)量表(表 3-3-1)由 Nasreddine 等 2004 年编制的用于快速筛查轻度认知功能损害(MCI)的评定工具。本节介绍的中文版是北京解放军总医院解恒革翻译的版本，因为这是国内最早翻译、验证并推广的版本。如果读者采用的是国内其他版本，只是测验材料(如要记忆的单词与要命名的动物)有差异，步骤与评分方法是相同的。

表 3-3-1　MoCA 量表记录纸

视空间/执行功能　　　　　　　　　　　　　　　　　　　　　　POINTS



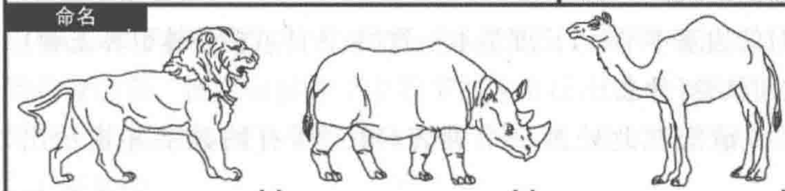
复制立方体



画钟（11点10分）（3分）

[ ]　　[ ]　　[ ] 轮廓　　[ ] 数字　　[ ] 指针　　__/5

命名



[ ]　　[ ]　　[ ]　　__/3

记忆		脸面	天鹅绒	教堂	雏菊	红色	
阅读名词清单，必须重复阅读。读2次 在5分钟后回忆一次	第1次						No points
	第2次						

注意力	现在我阅读一组数字（1个/秒）　顺背 [ ] 2 1 8 5 4　倒背 [ ] 7 4 2	__/2
	现在我阅读一组字母，每当读到A时请用手敲打一下。错2个或更多得0分　[ ] FBACMNAAJKLBAFAKDEAAAJAMOFAAB	__/1
	现在请您从100减去7，然后从所得的数目再减去7，共计算五次　[ ] 93　[ ] 86　[ ] 79　[ ] 72　[ ] 65　连减：4或5个正确得3分，2或3个正确得2分，1个正确得1分，0个正确得0分	__/3
语言	现在我说一句话，请清楚地重复一遍，这句话是："我只知道今天李明是帮过忙的人"。"当狗在房间里的时候，猫总是藏在沙发下"。	__/2
	流畅性/固定开头词语 "请您尽量多地说出以"发"字开头的词语或俗语，如"发财"。我给您1分钟时间，您说得越多越好、越快越好，尽量不要重复。" [ ] ____ (N ≥ 11 words)	__/1
抽象能力	请说出它们得相似性。　例如：香蕉 - 桔子　火车 - 自行车　手表 - 尺	__/2

	没有提示	脸面	天鹅绒	教堂	雏菊	红色		
		[ ]	[ ]	[ ]	[ ]	[ ]	Points for UNCUED recall only	__/5
选项	类别提示							
	多选提示							

定向力	[ ] 星期　[ ] 月份　[ ] 年　[ ] 日　[ ] 地方　[ ] 城市	__/6

www.mocatest.org　　　Normal ≥ 26 / 30　　TOTAL __/30　Add 1 point if ≤ 12 yr edu

二、MoCA 的使用说明

1. 交替连线测验

指导语："我们有时会用'123……'或者汉语的'甲乙丙……'来表示顺序。请您按照从数字到汉字并逐渐升高的顺序画一条连线。从这里开始[指向数字(1)]，从 1 连向甲，再连向 2，并一直连下去，到这里结束[指向汉字(戊)]。"

评分：当患者完全按照"1-甲-2-乙-3-丙-4-丁-5-戊"的顺序进行连线且没有任何交叉线时给 1 分。当患者出现任何错误而没有立刻自我纠正时，给 0 分。

2. 视空间技能(立方体)

指导语(检查者指着立方体)："请您照着这幅图在下面的空白处再画一遍，并尽可能精确。"

评分：完全符合下列标准时，给 1 分；只要违反其中任何一条，即为 0 分。

(1) 图形为立方体结构。

(2) 所有的线都存在。

(3) 无多余的线。

(4) 相对的边基本平行，长度基本一致(长方体或棱柱体也算正确)。

3. 视空间技能(钟表)

指导语："请您在此处画一个钟表，填上所有的数字并指示出 11 点 10 分。"

评分：符合下列 3 个标准时，分别给 1 分；如果违反其中任何一条，则该项目不给分。

(1) 轮廓(1 分)：表面必须是个圆，允许有轻微的缺陷(如，圆没有闭合)。

(2) 数字(1 分)：所有的数字必须完整且无多余的数字；数字顺序必须正确且在所属的象限内；可以是罗马数字；数字可以放在圆圈之外。

(3) 指针(1 分)：必须有两个指针且一起指向正确的时间；时针必须明显短于分针；指针的中心交点必须在表内且接近于钟表的中心。

4. 命名

指导语：自左向右指着图片问患者："请您告诉我这个动物的名字。"

评分：每答对一个给 1 分。正确回答是：① 狮子；② 犀牛；③ 骆驼或单峰骆驼。

5. 记忆

指导语：检查者以每秒钟1个词的速度读出5个词，并向患者说明："这是一个记忆力测验。在下面的时间里我会给您读几个词，您要注意听，一定要记住。当我读完后，把您记住的词告诉我。回答时想到哪个就说哪个，不必按照我读的顺序。"把患者回答正确的词在第一试的空栏中标出。当患者回答出所有的词，或者再也回忆不起来时，把这5个词再读一遍，并向患者说明："我把这些词再读一遍，努力去记并把您记住的词告诉我，包括您在第一次已经说过的词。"把患者回答正确的词在第二试的空栏中标出。

第二试结束后，告诉患者一会儿还要让他回忆这些词："在检查结束后，我会让您把这些词再回忆一次。"

评分：这两次回忆不记分。

6. 注意力

(1) 数字顺背广度

指导语："下面我说一些数字，您仔细听，当我说完时您就跟着照样背出来。"按照每秒钟1个数字的速度读出这5个数字。

评分：复述准确，每一个数列分别给1分。

(2) 数字倒背广度

指导语："下面我再说一些数字，您仔细听，但是当我说完时您必须按照原数倒着背出来。"按照每秒钟1个数字的速度读出这5个数字。

评分：复述准确，每一个数列分别给1分(注：倒背的正确回答是2、4、7)。

(3) 警觉性

指导语：检查者以每秒钟1个的速度读出数字串，并向患者说明："现在我朗读一组字母，每当我读到A时请用手敲打一下。其他的字母不要敲打。"测试员以1个/s的速度朗读字母序列。

评分：如果完全正确或只有一次错误则给1分，否则不给分(错误是指当读A的时候漏敲，或读其他字母时误敲)。

(4) 连续减7

指导语："现在请您做一道计算题，从100中减去一个7，而后从得数中再减去一个7，一直往下减，直到我让您停下为止。"如果需要，可以再向患者讲一遍。

评分：本条目总分3分。全部错误记0分，1个正确给1分，2到3个正确给2分，4到5个正确给3分。从100开始计算正确的减数，每一个减数都单

独评定，也就是说，如果患者减错了一次，而从这一个减数开始后续的减7都正确，则后续的正确减数要给分。例如，如果患者的回答是93－85－78－71－64，85是错误的，而其他的结果都正确，因此给3分。

7. 句子复述

指导语："现在我要对您说一句话，我说完后请您把我说的话尽可能原原本本地复述出来[暂停一会儿]：我只知道今天李明是来帮过忙的人。"患者回答完毕后，"现在我再说另一句话，我说完后请您也把它尽可能原原本本的复述出来[暂停一会儿]：狗在房间的时候，猫总是躲在沙发下面。"

评分：复述正确，每句话分别给1分。复述必须准确。注意复述时出现的省略(如，省略了"只"、"总是")以及替换/增加(如"我只知道今天李明……"说成"我只知道李明今天……"；或"房间"说成"房子"等)。

8. 词语流畅性

指导语："请您尽量多地说出以'发'字开头的词语或俗语，如'发财'。时间是1 min，您说得越多越好，越快越好，尽量不要重复。"

评分：在1 min内说出11个或者更多的词语则记1分。同时在空白处记下患者的回答内容。

9. 抽象　让患者解释每一对词语在什么方面相类似，或者说他们有什么共性。指导语从例词开始。

指导语："请您说说橘子和香蕉在什么方面相类似?"如果患者回答的是一种具体特征(如，都有皮，或都能吃等)，那么只能再提示一次："请再换一种说法，它们在什么方面相类似?"如果患者仍未给出准确回答(水果)，则说："您说的没错，也可以说它们都是水果。"但不要给出其他任何解释或说明。

在练习结束后，说："您再说说火车和自行车在什么方面相类似?"当患者回答完毕后，再进行下一组词："您再说说手表和尺子在什么方面相类似?"不要给出其他任何说明或启发。

评分：只对后两组词的回答进行评分。回答正确，每组词分别给1分。只有下列的回答被视为正确："火车和自行车：运输工具；交通工具；旅行用的。""手表和尺子：测量仪器；测量用的。"

下列回答不能给分："火车和自行车：都有轮子。""手表和尺子：都有数字。"

10. 延迟回忆

指导语："刚才我给您读了几个词让您记住，请您再尽量回忆一下，告诉

我这些词都有什么?”对未经提示而回忆正确的词，在下面的空栏中打钩(√)作标记。

评分：在未经提示下自由回忆正确的词，每词给1分。在延迟自由回忆之后，对于未能回忆起来的词，通过语义分类线索鼓励患者尽可能地回忆。经分类提示或多选提示回忆正确者，在相应的空栏中打钩(√)作标记。先进行分类提示，如果仍不能回忆起来，再进行多选提示，例如，“下列词语中哪一个是刚才记过的：鼻子、面孔、手掌?”各词的分类提示和(或)多选提示见表3-3-2。

表3-3-2 各词的分类提示和(或)多选提示

目标词语	类别提示	多选提示
面孔	身体的一部分	鼻子、面孔、手掌
天鹅绒	一种纺织品	棉布、的确良、天鹅绒
教堂	一座建筑	教堂、学校、医院
菊花	一种花	玫瑰、菊花、牡丹
红色	一种颜色	红色、蓝色、绿色

线索回忆不记分。线索回忆只用于临床目的，为检查者分析患者的记忆障碍类型提供进一步的信息。对于提取障碍导致的记忆缺陷，线索可提高回忆成绩；如果是编码障碍，则线索无助于提高回忆成绩。

11. 定向

指导语：“告诉我今天是什么日期。”如果患者回答不完整，则可以分别提示患者：“告诉我现在是[哪年，哪月，今天确切日期，星期几]。”然后再问：“告诉我这是什么地方，它在哪个城市?”

评分：每正确回答一项给1分。患者必须回答精确的日期和地点(医院、诊所、办公室的名称)。日期上多一天或少一天都算错误，不给分。

三、MoCA使用的注意事项

目前针对MoCA的划界分有颇为不同的报道。

一组的研究以25或26分作为MoCA划界分。编撰者Nasreddine通过对90例健康对照、94例MCI(其中aMCI-s 90例、aMCI-m 4例、没有naMCI)和93例轻度AD的MoCA检查，发现以MoCA总分≤26分作为划界分，敏感性90%，特异性87%。此后有3篇报道具有相似的结果：① Smith报道23例MCI、12例来自记忆门诊的正常对照组，采用同样的划界分，敏感性

83%，特异性为 50%，在随访 6 个月后，35%发展为痴呆，其基线 MoCA≤26 分。② 温洪波等采用 MoCA 评估 215 例健康老人、66 例 MCI(教育程度在初中以上)，以 MoCA 总分≤25 分作为划界分，敏感性 92.4%，特异性 88.4%。③ 张立秀等评估广州 230 例 60 岁以上老人(其中健康对照组 79 例、MCI 组 78 例、痴呆组 73 例)，在总样本中，文盲和小学占 58%，以 MoCA 总分≤24 分作为划界分，敏感性87.2%，特异性 83.6%。

另一组的研究以 22 分作为划界分。如：① 韩国 Jun-Young Lee 采用 MoCA－K(MoCA 韩国版)评估 196 例老人(其中健康对照组 115 例、MCI 组 37 例、轻度 AD 组 44 例)，以 MoCA 总分≤22 分作为划界分，敏感性 89%，特异性 84%。② Luis等调查美国 Sarasota 地区 74 例健康对照、24 例 MCI、20 例 AD 患者，当 MoCA 总分≤25 作为划界分时，识别 MCI 的敏感性为 97%，但是，特异性仅为 35%，≤22 分时，识别 MCI 的敏感性为 96%，特异性为 95%。③ 作者 2009 年调查教育程度在 5 年≤教育程度≤8 年、9 年≤教育程度≤12 年、教育程度≥13 年的老人中，MoCA 总分的划界分分别为≤21 分、22 分、23 分，得到的识别 MCI 的敏感性为 76.0%(其中识别aMCI－s 的敏感性 70%，识别 aMCI－m 的敏感性 93%，识别 naMCI 的敏感性 55%)，特异性为 80%。

贾建平等(2012)进行了基于社区的大样本调查，共8 411 例 65 岁及以上的社区老人，6 283 例认知功能正常，1687 例 MCI，441 例痴呆。划界分：教育程度为文盲者，13/14 分；1～6 年者 19/20 分；≥7 年者 24/25 分，敏感性 83.8%，特异性 80.5%。

很显然，MoCA 的另一个缺点是文盲和低教育老人中部分项目不适合，如模仿立方体和画钟对于没有书写经验的老人是不能完成的，连线和相似性的指导语也不容易为低教育老人所理解。虽然信息加工和反应速度是 naMCI 最敏感的指标之一，MoCA 的所有项目是不计时的，总耗时数的延长往往被忽视了。

与 MMSE 一样，MoCA 的分析指标是总分：把右侧栏目中各项得分相加即为总分，满分 30 分。量表设计者的英文原版应用结果表明，如果受教育年限≤12 年则加 1 分，最高分为 30 分。≥26 分属于正常。MoCA 已经在国内不少单位使用，常常采用项目分代替因子分的情况。事实上，项目分并不反映相应的认知领域，如 MoCA 的简短连线项目和相似性项目与完整的连线测验和相似性测验相关性极低。所以，MoCA 不能用于区分 MCI 亚型。MoCA 所

有项目不计时，而信息加工速度是执行功能（VCI－ND主要缺损领域）最敏感的指标。Sarah T. Pendlebury（2012）等以NINDS－CSN的VCI标准化神经心理测验为金标准诊断V－MCI，91例TIA或卒中患者，39例（42％）有MCI，其中多领域19例，单领域20例。MoCA的敏感性77％，特异性83％，ACE－R的敏感性83％，特异性73％。由于缺乏加工速度指标，这两种筛查测验对单领域受损MCI不敏感。

人进入老年期以后，躯体和精神疾病增多，轻度的视听力障碍、脑动脉硬化、情绪问题等相当普遍，从而对细致的敏感的认知功能检测结果产生一定的影响，增加了MCI识别和界定的难度，这也是影响各种筛查测验的敏感性和特异性的共同原因。

第四节 记忆与执行筛查量表（MES）

一个理想的针对MCI的认知筛查工具应简明，耗时短，易于执行，不受教育、文化、语言等混杂因素影响，能够反映其最突出最主要的认知损害知域。目前国际上还没有公认的符合以上要求的MCI筛查工具。

记忆与执行筛查量表（memory and executive screening，MES）是作者2009年编制的MCI的筛查量表（表3－4－1）。MES满分100分，记忆因子分50分，执行因子分50分。

表3－4－1 记忆与执行筛查量表（MES）记录纸

问题1（问受试者）：如果你以往的记忆力表现可以打100分，那么，现在你可以给自己大约打多少分？

问题2（问家属）：如果老人以往的记忆力表现可以打100分，那么，你认为他（她）现在大约可以打多少分？

序号	项目	满分
1	N1[李][小明]有[2只][灰色]的[小狗]，住在[永安][县][河西][镇][58号] N2[李][小明]有[2只][灰色]的[小狗]，住在[永安][县][河西][镇][58号] N3[李][小明]有[2只][灰色]的[小狗]，住在[永安][县][河西][镇][58号]	10 10 10
2	流畅性[列举“厨房里有的东西”，30秒，全部记录]：	10
3	扣指：(1) 矛盾刺激“我敲二次，你敲一次。我敲一次，你敲二次。” 1－1－2－1－2－2－2－1－1－2－1－2－2－1－1	10
4	短延迟回忆：第4次回忆（不再复述）： [李][小明]有[2只][灰色]的[小狗]，住在[永安][县][河西][镇][58号]	10

（续表）

序号	项　　　目	满分
5	手指结构：① 拇指连续接触另 4 指的指尖[右　　左　　] ② 拇指夹在示指和中指间——剪刀状[右　　左　　] ③ 手卷起来放在眼睛前——同侧耳朵上——嘴前[右　　左　　] ④ 画十字[右　　左　　] ⑤ 攥拳—掌缘—手平放[右　　左　　]	20
6	扣指：(2) 敲一不敲“我敲一次，你也敲一次。我敲二次，你不敲。” 1-2-1-2-1-1-2-2-1-1-2-1-2-1-2	10
7	长延迟回忆：第 5 次回忆 [李][小明]有[2 只][灰色]的[小狗]，住在[永安][县][河西][镇][58 号]	10

作者调查 197 例正常对照组老人、116 例 aMCI - sd 患者，195 例 aMCI - md 患者，228 例轻度 AD 患者(来源于 2009 年 6 月～2011 年 10 月华山医院记忆门诊)。全部样本在 50 岁以上，各组之间的年龄、性别、教育没有显著差异。评估量表除了 MMSE、MES，还有 AVLT、TMT、BNT 等标准化神经心理测验，并以这些标准化神经心理测验得分作为 MCI 诊断的依据。

相关分析发现 MES 的 3 个指标(记忆得分、执行得分和总分)与年龄有显著相关性，年龄愈大得分愈低，与教育程度没有显著相关性。MES 没有明显的天花板效应和地板效应。完成 MES 的耗时数平均约 7 min。ROC 分析表明 MES 总分≤75 分识别 aMCI - sd 组的曲线下面积为 0.89，敏感性 80%，特异性 83%，MES 总分≤72 分识别 aMCI - md 组的曲线下面积为 0.95，敏感性 87%，特异性 91%。

MES 具有比较高的敏感性和特异性、易操作、易携带(不需要任何辅助材料)、耗时少(少于 MMSE 耗时数)，适合用于门诊或社区的 MCI 的筛查。期待有兴趣的研究者将 MES 与 MoCA 进行比较研究，并进行长期随访，不仅验证其早期诊断价值，也验证其治疗效果评估价值。MES 的使用说明见表 3 - 4 - 2。

表 3 - 4 - 2　记忆与执行筛查量表(MES)的使用说明

第 1 题：施测者读出以下一句话，不要给受试者看 “李小明有 2 只灰色的小狗，住在永安县河西镇 58 号。”	注意：① 要一口气读完，不要在中途回答受试者的提问 ② 第三次回忆结束时，告知受试者“请把它记住，等会要请你回忆。” ③ 假如受试者第一或第二次就全部回忆正确，仍然需要学习第三次 评分方法：共 10 个要点。每个要点完全正确得 1 分，部分正确(如 58 号说成 38 号)也不得分，“灰色的小狗”说成“小灰狗”也不得分。只要说出要点，非要点(如“有”、“住在”)不管是否回答不记分。要点的次序颠倒(如说成“李小明的 2 只小狗是灰色的”)不扣分

（续表）

第2题：指导语："请你列举尽可能多的厨房用的或看到的东西，愈多愈好"，30 s结束，分时记录全部回答，超过10个也要记录下来	评分方法：正确列举1个1分，满分10分
第3题：扣指测验——矛盾刺激 "我敲二次，你敲一次。我敲一次，你敲二次。"为了使患者理解指导语，举例：检查者敲1次，患者跟着敲2次；检查者敲2次，患者跟着敲1次。每次扣指约1 s	评分方法：满分10分。圈出错误数字，每个错误扣除1分，直到0分。记录耗时数 注意： ① 受试者理解了才能开始做，正式开始后以均速进行，会打断，说完为止 ② 避免给受试者应该敲还是不敲的暗示 ③ 受试者扣指以敲到桌面为准，到半途停止的不计 ④ 为了避免敲伤受试者手指，敲桌子也可以是拍桌子，或手握铅笔敲桌子
第4题：指导语："现在请你回忆开头学习过的那句话。"	评分方法与第1题相同
第5题：手指结构测验 指导语："我现在做一些动作，你跟着我做。"不要用语言描述这些动作。 操作步骤： 第1步：利手模仿动作，单手完成； 第2步：对于不能正确完成的动作，可以重复模仿一遍； 第3步：施测者没有进行非利手示范的情况下，受试者用非利手操作一遍； 第4步：如果受试者非利手不能正确完成，施测者可以用非利手重复一遍。 每个动作最多模仿2次	评分方法：都是连续动作，要求动作次序和手势形态同时正确（即时间和空间两方面同时正确）。只对正确的动作评分：2分＝第一次动作正确，1分＝第二次动作正确，0分＝2次动作都不正确 动作： ① 拇指连续接触示指、中指、环指、小指的指尖（连续组成4个0形） ② 拇指夹在屈曲的示指和中指间，其他手指屈曲——剪刀状（示指和中指作剪刀的动作，其他手指屈曲） ③ 手像试管一样卷起来放在眼睛前（望远镜状）——同侧耳朵上（倾听状）——嘴巴前面（喝水状） ④ 画十字（依次触碰前额、前胸、对侧和同侧的肩膀） ⑤ 攥拳头—掌缘往下砍——手指并拢平放在桌面、手背朝下
第6题：扣指测验——敲-不敲 "我敲一次，你也敲一次。我敲二次，你不敲。"为了使患者理解指导语，举例：检查者敲1次，患者敲1次。检查者敲2次，患者不敲。值到受试者表示理解才开始做，进行过程中有错误不予提醒。每次扣指约1 s	评分方法：满分10分。圈出错误数字，每个错误扣除1分，直到0分。记录耗时数
第7题：指导语："现在请你再次回忆开头学习过的那句话。"	评分方法与第1题相同

（郭起浩）

第五节　蒙特利尔认知评估基础量表(MoCA-B)

自从蒙特利尔认知评估量表(MoCA)2005 年正式发表以来,国内外已经有大量研究证实 MoCA 在识别 MCI 方面优于 MMSE、剑桥认知评估(CAMCOG)、画钟测验(CDT)、言语流畅性测验(列举动物或水果)、功能活动问卷(Pfeffer FAQ)等常用筛查量表。

在生态效度方面,也积累了不少资料,如 Del Brutto(2014)调查了 311 名厄瓜多尔社区老人[平均年龄 69.2 岁,199 (83%)小学文化],平均 MoCA 得分 18.5 分,在控制各种潜在影响因素后,发现 MOCA 总分、主要领域分与中重度患者的总体皮质萎缩(GCA)程度显著相关。Hollis(2015)调查出事故的驾车者的认知功能,发现是 MoCA 而不是 MMSE 可以预测驾车安全状况。

然而,MoCA 在发展中国家应用有一个普遍存在的问题,就是文盲与低教育老人的适用性。有许多研究将低教育者人为地加 1 分或 2 分,因为保持原来的划界分,显然不能解决文盲老人的评估问题。唐牟尼教授发表在《中华精神科杂志》(2014)上,针对广州市老城区、新城区、郊区分层整群抽样的2 111名老人的流行病学调查发现,文盲组人口占 40%,小学组人口占 38%,中学及以上组人口只占 22%,作者的结论之一是文盲组 MCI 的识别,MMSE 优于 MoCA,另两组无显著差异。王华丽教授发表在《科学世界杂志》(*Scientific World Journal*)上的小样本研究得到类似结果。贾建平教授 2012 年发表的常模,用于识别 MCI 的划界分,发现社区老人教育程度≥7 年者 24/25分、1～6 年者 19/20 分、文盲者 13/14 分;显然不是增加 1～2 分就可以调整过来的。

为了适于低教育老人的认知评估,Nasreddine 于 2014 年 6 月发表了新版本,称为蒙特利尔认知评估基础量表(Montreal cognitive assessment-basic, MoCA-B),并在泰国进行了初步的信度与效度验证。笔者于 2015 年 7 月与 Nasreddine 取得联系,签署了中文版翻译协议,经过大量讨论与回译,最终确定了中文版本,目前该版本已经可以在 MoCA 官方网站(www.mocatest.org)上自由下载了。

单纯从项目选择与组成(表 3-5-1)上看,MoCA-B 弥补了 MoCA 的许多争议点,如删除了执笔项目画立方体与画钟,替换了动物命名与相似性题

目，将流畅性与划销项目进行分级评分，计算题目更贴近生活，删除了顺背、倒背与复述这些对施测者要求比较高的项目，由 12 组题减为 9 组题，从而节省评估时间。

表 3-5-1 MoCA-B 与 MoCA 在项目与评分方面的差异

	项 目	MoCA 项目	得分	MoCA-B 项目	得分
1	视空间/执行功能	连线 1	1	连线 1	1
		画立方体 1、画钟 3	4	视知觉 3	3
2	命名	动物 3 种	3	动物 4 种	4
3	注意力	顺背 1、倒背 1、叩击 1	3	朗读数字	3
		100 连续减 7	3	计算力	3
4	语言	复述 2、流畅性 1	3	流畅性 2	2
5	抽象思维能力	相似性 2 题	2	相似性 3 题	3
6	延迟回忆		5		5
7	定向力		6		6

一、测试指导和评分方法

MoCA 是一个用来对轻度认知功能损害(MCI)进行快速筛查的评定工具。量表已在高教育程度的老年人(平均教育年限 13 年)中验证其发现 MCI 患者及鉴别患者与健康老年人的能力。但量表中许多项目受教育程度的影响较大。故笔者设计了用于筛查文盲和低教育程度人群 MCI 的新版 MoCA，即 MoCA-B(表 3-5-2、表 3-5-3)。MoCA-B 与原版 MoCA 评估相同的认知领域：执行功能、语言、定向、计算、抽象思维、记忆、视知觉(而不是视结构技能)、注意和集中。MoCA-B 的测试时间约 15 min，总分 30 分。

开始时间：在开始给受试者介绍第一部分测试(执行功能)时开始计算时间(时-分-秒)，记录于量表的右上角。

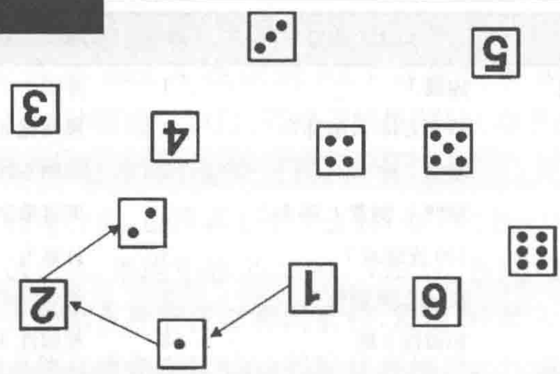
1. 执行功能(交替连线测验)

指导语：检查者向受试者说明："请您按照从数字到点并逐渐升高的顺序画一条连线。从这里开始[指向数字(1)]，从数字 1 连向一个点[指向含有一个点的正方形]，再连向数字 2[指向数字(2)]，之后连向两个点[指向含有两个点的正方形]，并一直连下去，到这里结束[指向含有 6 个点的正方形]。"

评分：当受试者完全按照顺序进行连线时给 1 分。当受试者未按顺序连线或出现任何错误时，给 0 分。

表 3-5-2 蒙特利尔认知评估基础量表中文版

姓名______
性别______ 年龄______
教育年限______ 测试日期______
检查者______

执行功能	得分
	开始时间 ______ (/1)

即刻回忆		桃花	萝卜	沙发	蓝色	筷子	不计分
即时第一次测试所有词语均能回忆，也需完成第二次测试。	第一次						
	第二次						

流畅性	在 1 分钟内尽可能多的说出水果的名字			n= 个	(/2)
1~15 秒：	16~30 秒：	31~45 秒：	46~60 秒：	n≥13 计 2 分 n=8~12 计 1 分 n≤7 计 0 分	

定向	[ ]时间(±2 小时) [ ]星期几 [ ]月份 [ ]年份 [ ]地点 [ ]城市	(/6)

计算 用 1 元、5 元、10 元钱购买"13 元"的物品，说出 3 种付款方式。（说出 3 种正确付款方式计 3 分，2 种计 2 分，1 种计 1 分，未说出计 0 分） (/3)

正确方式：①②③④错误方式：

抽象 下面的事物属于什么类别？（例如：香蕉-橘子=水果） (/3)

[ ]火车-轮船 [ ] 锣鼓-笛子 [ ] 北方-南方

延迟回忆	回忆时不提示	桃花	萝卜	沙发	蓝色	筷子	(/5)
		[ ]	[ ]	[ ]	[ ]	[ ]	
未经提示下自由回忆正确的词计分（每词 1 分）	分类提示	[ ]	[ ]	[ ]	[ ]	[ ]	
	多选提示	[ ]	[ ]	[ ]	[ ]	[ ]	

视知觉 图片识别，时间 60 秒。图片见附页。

剪刀	T 恤	香蕉	台灯	蜡烛
手表	杯子	叶子	钥匙	勺子

n=9~10 计 3 分
n=6~8 计 2 分
n=4~5 计1分
n=0~3 计 0 分 (n=) (/3)

命名 动物命名，图片见附页。[ ]斑马 [ ]孔雀 [ ]老虎 [ ]蝴蝶 (/4)

注意 朗读**圆形**中的数字：数列见附页 158392039402168746755 5 错误数__n 错误数≤1 个计 1 分 (/1)

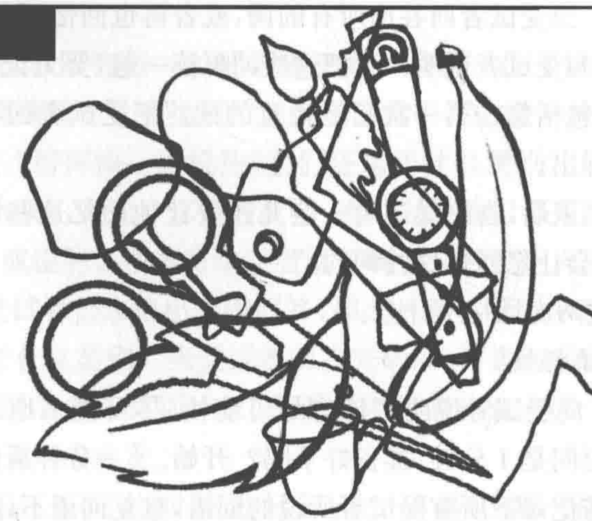
朗读**圆形**和**正方形**中的数字：38513029204978615764 数列见附页 158392039402168746 75 错误数__n 错误数≤2 计 2 分 错误数=3 计 1 分 错误数≥4 计 0 分 (/2) 结束时间 ______

Adapted by:Qihao GuoMDChinese versionAugust 01,2015

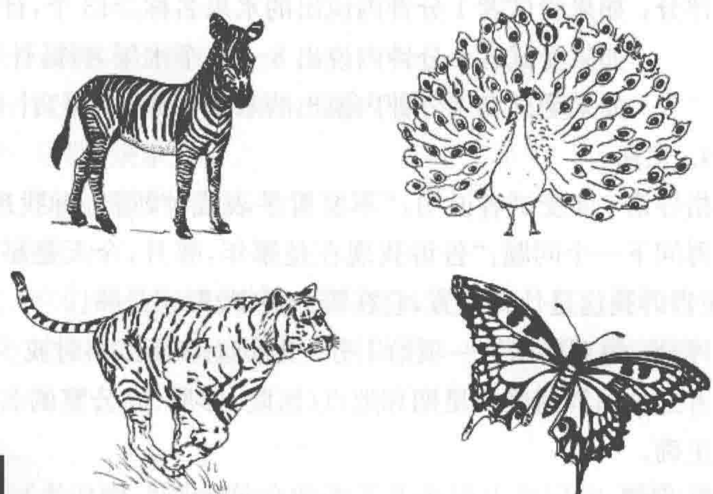
总分 (/30)
受教育年限<4 年加 1 分，不识字再加 1 分
总时间 分 秒

表 3-5-3 蒙特利尔认知评估基础量表中文版(附页)

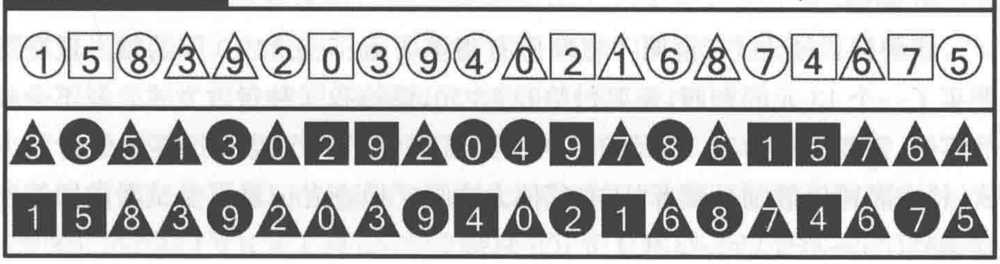
视知觉



命名



注意



2. 即刻回忆

指导语：检查者向受试者说明："这是一个记忆力测验。下面我会给您读5个词，您要注意听，一定要记住。当我读完后，把您记住的词告诉我。回答时想到哪个就说哪个，不必按照我读的顺序。"检查者以每秒钟1个词的速度读

出 5 个词(桃花、萝卜、沙发、蓝色、筷子)。把受试者回答正确的词在第一试的空栏中标出。当受试者回答出所有的词,或者再也回忆不起来时,把这 5 个词再读一遍,并向受试者说明:“我把这些词再读一遍,努力记住它们并把您记住的词告诉我,包括您在第一次已经说过的词。”把受试者回答正确的词在第二试的空栏中标出。

第二试结束后,告诉受试者一会儿还要让他回忆这些词:“请您记住这些词,我之后还会让您回忆这些词的。”

评分:这两次回忆不计分。

3. 词语流畅性

指导语:向受试者说明:“请您尽可能快、尽可能多地说出您所知道的水果的名称。时间是 1 分钟,准备好了吗? 开始。[一分钟后停止]结束。”

检查者需记录下所有受试者所说的词语,重复词语不计入得分。

评分:如果受试者 1 分钟内说出的水果名称≥13 个,计 2 分。

如果受试者 1 分钟内说出 8~12 个水果名称,计 1 分。

如果受试者 1 分钟内说出的水果名称≤7 个,计 0 分。

4. 定向

指导语:向受试者说明:“不要看手表或钟,请告诉我现在是几点钟了?”然后再问下一个问题:“告诉我现在是哪年,哪月,今天是星期几?”最后再问:“现在告诉我这是什么地方,它在哪个城市?”

评分:每正确回答一项给 1 分。时间上多于 2 小时或少于 2 小时都正确。受试者必须回答精确的星期和地点(医院、诊所、办公室的名称)。回答当地年月也正确。

5. 计算

指导语:向受试者说明:“想象您有很多 1 元、5 元和 10 元的钱。现在您购买了一个 13 元的东西,需要付给我 13 元,请给我 3 种付款方式。我不会找您零钱,需要您付给我 13 元整。”当受试者提供了一个需要找零钱的付款方式,检查者可以鼓励受试者“还有其他方法吗?”检查者记录下受试者的回答所指编号(① 一张 10 元+3 张 1 元;② 两张 5 元+3 张 1 元;③ 1 张 5 元+8 张 1 元;④ 13 张 1 元)。

评分:如果受试者提供 3 种正确付款方式,计 3 分。

如果受试者提供 2 种正确付款方式,计 2 分。

如果受试者提供 1 种正确付款方式,计 1 分。

如果受试者未提供正确付款方式,计 0 分。

6. 抽象

指导语:让受试者回答每一对词语属于哪一类别。指导语从例词开始。“请您说说橘子和香蕉属于什么类别?”如果受试者回答的是一种具体特征,那么只能再提示一次:“请再换一种说法,它们还属于什么类别?”如果受试者仍未给出准确回答(水果),则说:“您说的没错,也可以说他们都是水果。”但不要给出其他任何解释或说明。在练习结束后说:“现在您再说说火车和轮船属于什么类别?” 如果受试者仍未给出准确回答,那么只能再提示一次:“请再换一种说法,它们还属于什么类别?”当受试者回答完毕后,再进行后面两组词:“您再说说锣鼓和笛子属于什么类别?”和“您再说说北方和南方属于什么类别?”不要给出其他任何说明或启发。

评分:只对后三组词的回答进行评分。回答正确,每组词分别给 1 分。

- 只有下列回答被视为正确:

a 火车和轮船:交通工具,旅行用的,运输工具,客运工具。

b 锣鼓和笛子:乐器,娱乐工具。

c 北方和南方:方向,地方,地点,地理位置。

- 下列回答不能给分:

a 火车和轮船:它们都是钢铁做的。它们都有发动机。它们都耗汽油。

b 锣鼓和笛子:它们都是木头或其他材料做的。它们都可以发声音。

c 北方和南方:地理。

7. 延迟回忆

指导语:向受试者说明:“刚才我给您读了几个词让您记住,请您再尽量回忆说出这些词。如果您不记得所有词语和它们的顺序,也不需要紧张。”对未经提示而回忆正确的词,在下面的空栏中打钩(√)做标记。

评分:在未经提示下自由回忆正确的词,每词给 1 分。

线索回忆指导语:在延迟自由回忆之后,对于未能回忆起来的词,通过语义分类线索鼓励受试者尽可能地回忆。经分类提示或多选提示回忆正确者,在相应的空栏中打钩(√)做标记。对所有未能回忆的词进行线索回忆。先进行分类提示,如果仍不能回忆起来,再进行多选提示。例如:“下列词语中哪一个是刚才记过的:桃花、菊花、梅花?”

各词的分类提示和(或)多选提示如下:

分类提示	多选提示
桃花：一种花	**桃花**、梅花、菊花
萝卜：一种蔬菜	南瓜、洋葱、**萝卜**
沙发：一种家具	桌子、**沙发**、椅子
蓝色：一种颜色	**蓝色**、绿色、红色
筷子：一种厨房用具	刀子、勺子、**筷子**

附件 3－5：例词的修改

原文翻译为：

玫瑰：一种花	玫瑰、雏菊、郁金香
椅子：一种家具	桌子、椅子、床
手：身体的一部分	脚、手、膝盖
蓝色：一种颜色	蓝色、棕色、红色
勺子：一种厨房用具	叉子、刀子、勺子

修改理由：

(1)“玫瑰”与“郁金香”在西方是常用的花朵名称，但译成中文后让受试者记忆比较难，因为都不是中国常见的花朵，所以改为常见的容易理解的频度相近的梅花、桃花。

(2)“椅子”通常刚好出现在施测者坐的地方，被试时容易形成视觉记忆，已经有大量研究证实听觉记忆识别 MCI 的敏感性与特异性明显优于视觉记忆，所以，笔者改为同样常见但测试现场看不到的名词“沙发”(Guo Q. ADAD. 2009;23(3)：253－259. Kim MJ, et al. ADAD. 2011, 25(3)：242－249. Zhao Q, et al. CAR. 12(6)：520－526)。

(3)“手”与“椅子”一样，就在受试者眼前，为了避免视觉记忆，笔者修改为频度相近的名词“萝卜”。

(4) 蓝色，保持不变。多选部分，将棕色改为容易被理解的绿色。

(5)“spoon”的中文意思是“勺子”，容易出现歧义。在中国的许多地方，勺子指代不同的东西，可以是吃饭用的调羹、锅铲，也可以是木工用的工具，所以，改为没有歧义的常见的“筷子”。

以上修改理由已翻译为英文并与 Nasreddine 沟通，获得同意并受好评，也经过多个不同方言地区的试测。笔者希望 MoCA－B 能够避免 MoCA 在中国出现的混乱状况(MoCA 至少有 6 个中文版本，常常是同一单位不同研究小组

采用的版本都不一样），以至于即使国内不同单位的数据亦无法直接比较、无法开展meta分析的现状。

线索回忆评分：不计分。线索回忆只用于临床目的，为施测者分析记忆障碍类型提供进一步的信息。对于提取障碍导致的记忆缺陷，线索可提高回忆成绩；如果是编码障碍，则线索无助于提高回忆成绩。

8. 视知觉

指导语：检查者指向附录中视知觉图片，并告诉受试者："现在请您看这张图。图片里有很多重叠在一起的物品。请尽可能地把它们找出来。如果您不知道它们的名字，可以指出它们的轮廓或告诉我它们的功能。不能旋转图片。你可以慢慢做，但时间不超过2 min。准备好了吗？开始。"

指导语结束60 s后停止测试。受试者不能旋转图片，施测者不能告知受试者总共有10项物品。在视知觉部分计分表上用数字记录每个正确回答的顺序。

评分：图片中有10个物品：剪刀、杯子、T恤（衬衣、内衣）、手表、香蕉、叶子（树叶）、台灯、钥匙（锁匙）、蜡烛和调羹（勺子）。

如果受试者找出9～10个物品，计3分。

如果受试者找出6～8个物品，计2分。

如果受试者找出4～5个物品，计1分。

如果受试者找出3个或3个以下物品，计0分。

9. 命名

指导语：自左向右、从上到下指着附录中的图片问受试者："请您告诉我这个动物的名字。"

评分：每答对一个给1分。正确回答是：① 斑马[马和驴不得分]。② 孔雀[鸟不得分]。③ 老虎[猎豹、美洲豹和黑虎不得分]。④ 蝴蝶[昆虫不得分]。

10. 注意

指导语1：指向附录中白色背景的数字，并向受试者说明："请看这些白色背景的数字。现在要您大声读出圆形中的数字，正方形和三角形中的数字不要读。从这里开始[指向数列开头①]，到这里结束[指向数列结尾⑤]。开始。"

评分1：如果完全正确或只有一次错误，计1分。

如果有2个或2个以上错误，计0分。

错误是指读非圆形中的数字、跳过圆形中的数字而没有读、朗读数字顺序

错误或读之前的数字。记录下错误个数。

指导语 2：指向附录中黑色背景的数字，并向受试者说明："请看向这些黑色背景的数字。现在要您大声读出圆形和正方形中的数字，三角形中的数字不要读。从这里开始[指向第一行数列开头]，到这里结束[指向第二行数列结尾]。开始。"

评分 2：如果有 2 个或 2 个以下错误，计 2 分。

如果有 3 个错误，计 1 分。

如果有 4 个或 4 个以上错误，计 0 分。

错误是指读非圆形或正方形中的数字、跳过圆形或正方形中的数字而没有读、朗读数字顺序错误或读之前的数字。记录下错误个数。

结束时间：在受试者完成最后一项测试（注意）时停止计算时间（时-分-秒），计算测试时间（分钟，秒），记录于量表的右下角。

二、附加条件

除非有特殊要求，每个项目的测试指导语只能重复一遍。

总分：把右侧栏目中各项得分相加即为总分，满分 30 分。

为校正教育程度所致偏移，如果受试者受教育年限≤4 年则加 1 分，最高分为 30 分。如果受试者不识字，无论其受教育年限为多少，总分额外增加 1 分，最高分为 30 分。不识字定义为在日常生活中无法流利地读或写。

（郭起浩　陈科良）

第六节　非语言智力测验（TONI）

非语言智力测验（the test of nonverbal intelligence，TONI）是一个非语言智力测验。这个测验可在 6～89 岁 11 个月的人群中实施。Brown 等于 1982 发表第一版，1990 年发表修订版（TONI－2），1997 年第三次修订版面世、2009 年发表 TONI－4。TONI－2 比 TONI－1 每种形式增加了 5 个项目，TONI－3 比 TONI－2 更简短（每种形式减少了 10 个项目），而且有一个更新、更大的标准化样本。TONI－4 比 TONI－2 每种形式减少了 10 个项目，但增加了 5 个项目，每种形式 60 个项目。换句话说，四个版本中 85％以上的项目是一样的。

TONI是一个智力测验,不涉及实施过程中的语言、项目内容和反应形式。根据编制者的介绍,这使该测验非常适合于评估由于发育障碍或神经系统损伤导致失语、阅读障碍、语言障碍、学习障碍和语音障碍的患者。据报道,它也适用于评估英语能力有限而导致阅读或书写英语困难的人。没有文字、符号、数字和图画减少了这个测验的文化限制。

TONI的测验项目类似于矩阵模拟测试,每一页画有一个刺激图案和一组可能的反应(4个或6个选择)。评定员发出指令手势,受试者做出指向、点头或其他动作反应。有严重运动受损的患者可通过眨眼、止痛棒、光束或其他方法做出反应。

根据编制者的描述,TONI没有以任何智力理论为基础。但是,它是一个很强大的一般智力测验和液态智力测验,而不是晶态智力测验。与其他的非语言技能不同,TONI以抽象的推理和问题解决能力为基础,因为它们被认为是一般智力的主要组成部分,每天的生活都需要,可以形成一个不受语言限制和运动障碍影响的测试形式,是一个允许新奇而不受文化限制的项目。项目内容包括七个不同的抽象推理和问题解决能力:概括/分类、区别、类比、系列化、归纳、推论和细节识别。

Jensen(1980)建议,不受语言和文化限制的测验应该遵循7个原则:① 评估表现而不是书写能力;② 提供手势指令而不是口头或书面指令;③ 包含练习项目;④ 不计时的;⑤ 有抽象内容;⑥ 评估推理/问题解决能力而不是信息获得能力;⑦ 用新的问题,而不是之前学会的信息。根据作者所说,TONI满足以上所有的准则。

实施时间只要15~20 min,相比韦氏智力测验(WAIS-III)的80 min,可以说是智力测验领域的快速检测法。

一、实施

所有的受试者从第一项开始,当5个项目有3个答错时测试停止。测试项目以螺旋装订的书籍形式呈现。测试反应由评定员记在单独的记录单上,而且不计时。有两个等价的版本(A和B)。

二、评分

最初的得分需转换为离差智商(比如,标准得分格式的平均数为100,标准差为15),以及百分等级。只有一个换算表,而且两种形式都用,因为两种

形式是“心理测量等价的”。年龄当量(age equivalents)也能获得。然而在使用这些当量时要格外小心,因为年龄当量受众所周知的心理测量限制,也因为TONI年龄当量包括来自原始分数分布的平滑化(smoothing)、插入和推断。

三、人口统计学影响因素

1. 年龄　TONI-4得分与年龄(<17岁)密切相关,然后到达稳定,大约60岁开始下降。

2. 性别　手册表明女性的TONI-3 A版平均得分101,男性99。然而,手册并没有报道有意义的性别影响。

3. 教育　没有关于教育影响的信息。

4. 种族　手册表明以英语为第二语言的人、非裔美国人和西班牙裔人的TONI-3平均得分分别是93、95、96。这是该测试跨文化和种族适用性的证据。然而,就像DeMauro(2001)指出的,这意味着非裔美国人的得分大约比总样本平均分低1/3个标准差。这表明该测验可能不像作者所说的那样是不受文化限制的,可能会受文化的影响,像已存在的大多数所谓的非语言和不受文化影响的测验一样。

四、信度

1. 内部信度　内部信度很高。手册报道,标准化样本和年龄、种族、功能障碍等限定的选择性亚群的内部信度都很高。它们的系数α值为0.89~0.9。

2. 测量标准误　两种形式的平均测量标准误在3~5分。

3. 重测信度　重测信度也很高。针对170多名受试者,在间隔1个星期后再次接受了测验。3个年龄群的测验相关系数为0.89~0.94(13岁、15岁、19~40岁)。然而,没有<13岁和>40岁的受试者的稳定性信息。

4. 练习效应　手册中并没有关于练习效应的信息。然而,一个研究发现虽然学生合作获得的解决方法会使TONI-3得分提高,但是当学生再测试并被要求独立解决时,这个效应并没有继续存在。

5. 复本信度　手册报道,A版和B版的信度由立即实施范式测得(比如,两个版本都在一个测验过程中实施)。两个版本跨年龄的平均相关性是0.84。但是,两个版本间的相关性在所有的年龄群中高度一致(r为0.74~0.95),这意味着两个版本不一定在所有的年龄水平都是可以被替换的。像以上提及的

一样，一套标准测验的两个版本都要被用到。

6. 评定员间信度　评分者间信度很高(相关系数 $r = 0.99$)。

五、效度

1. 内容效度　在创建这个测验时，TONI 的项目都是谨慎选择的，非语言智力功能都是用经典的项目分析和差异功能分析来评估的，差异功能分析用来排除偏倚。

2. 因素结构　因素分析支持普遍因素的存在。

3. 与其他非语言智力测验之间的相关性　与其他非语言智力测验之间的相关性很高。比如，在一个550名的成人样本中，TONI-3与一个非语言智力测验——非语言智力综合测验的相关系数是0.76，与全量表智商得分的相关系数是0.74，与图画几何非语言智力量表的相关系数是0.64。最近希腊针对帕金森病患者的研究报道说，TONI-2和瑞文推理测验的相关系数是0.73。

4. 与标准智力测验的相关性　与标准智力测验的相关性很高。在一组学习障碍的学生中，TONI-3和韦氏儿童智力量表-Ⅲ全量表智商的相关性很高($r = 0.63$)，与言语智商和操作智商的相关性没有明显差别(r 分别为0.59，0.56)。在患有学习障碍的人群中，与 WAIS-R 的相关性也很高($r=0.73$)，与 PIQ 的相关性略高于 VIQ(r 分别为0.75，0.57)。要注意到这些相关性都是基于很小的样本。另外，没有关于这些测验平均表现差异的信息。

5. 与成就测验的相关性　与成就测验的相关性很高。比如，与 Woodcock-Johnson 心理—教育测验修订版的阅读与数学的相关性很高(r 分别为0.73，0.76)。值得注意的是，尽管认为 TONI-3 是评估新问题解决能力，而不是知识获得能力，它与 Woodcock-Johnson 心理—教育测验修订版知识分测验的相关性也很高($r=0.56$)。从使用 WRAT 和 WRAT-R 的两个研究中总结出，TONI 得分与阅读、数学和拼写有轻到中度的相关性($r=0.28$～0.38)。

六、临床研究

手册介绍了几个不同临床组的平均得分(比如聪明的失读患者、聋人、多动症儿童、情绪障碍和心理障碍者)。根据这些信息，有心理障碍的个体的平均得分最低。然而，并没有提供大量的临床相关性信息。各组的人口统计学

特点和诊断标准也没被描述。

根据手册呈现的一组聪明儿童的平均表现，Sattler(2001)指出 TONI－3 分辨出聪明个体(平均得分为 100)的能力有限。心理障碍患者的平均得分近似于诊断临界值 70 分，说明大多数智力缺陷者的得分高于临界值。

七、评价

TONI－4 的优点是简洁、不依赖于语言指令和反应、不依赖肢体活动，适用于绝大部分临床人群，是“用户友好”型测验。与许多测验不同的是，它还包括了关于不同亚组内部信度的信息。最后，它的选择模式是强迫性的，使受试者在没有最佳答案时会考虑次佳答案。然而，这个反应模式并没有被充分研究。

TONI 有它的不足之处，包括它的图形推理作为衡量智力的依赖性和评估天才的适用性有限。TONI 受地板效应的影响很大。这个测验不应该被用来诊断儿童或老年人的智力损害，也不能用来评估任何年龄阶段智力损害的严重程度(比如轻度、中度、重度)，这显著限制了它的临床应用。还有的担忧是它对手册中呈现的作为效度证据的临床组数据的覆盖和分析不足。没有足够的研究证据支持作者所说的“TONI 对于不同种族亚群(比如常说英语、不说英语以及有各种感觉和智力障碍的个体)是一个具有相同信度和效度的测量”。TONI 还有许多问题，如复本信度不够、不同版本间的相关性也不够充分、与语言能力测验和获得知识能力测验的显著相关性也没有被很好地解释。

虽然 TONI－4 作为一个筛查评估是有用的，但不能代替韦氏智力测验(WAIS－III)。然而，WAIS－III 由于版权、需要严格培训、耗时长等问题而不能普及到基层医院，而 TONI 避免了这些问题。笔者于 2014 年引进 TONI－4，但是在老年人中试用，难度明显偏高，是否可以在成人中应用，便于今后评估脑外伤、癫痫、脑炎后遗症等成人患者的智力改变，正在研究中。

(徐　岩　郭起浩)

第四章
记忆的评估

记忆是编码、储存和提取信息的复杂过程。编码是获得信息并加以处理和组合;储存是将组合整理过的信息做永久记录;检索或提取是将被储存的信息取出,回应一些暗示和事件。记忆可以分为各种成分,参见第一章第五节。常用的记忆评估量表包括 Fuld 物品记忆测验、各种版本的词语学习测验、Rey-Osterrieth 复杂图形测验(CFT)回忆部分、自传记忆晤谈(autobiographical memory interview,AMI)、Benton 视觉保持测验(BVRT)、再认记忆测验(recognition memory test,RMT)、Rivermead 行为记忆测验(RBMT)和各种版本的韦氏记忆测验。本节以介绍各种版本的词语学习测验为主,不像韦氏记忆测验有版权要求,这些词语学习测验通常可以自由地被使用,而且耗时较短、分析指标更全面。

第一节 Rey 听觉词语学习测验(RAVLT)

Rey 听觉词语学习测验(Rey auditory verbal learning test,RAVLT)也直接被称为 AVLT。其最早的版本是一个词表一次学习后回忆,发表于 1900 年代,1958 年由 Claparede 的学生 Rey 发表的修订版,需要 5 次学习过程。目前的版本还经过 Taylor(1959)和 Lezak(1976)的深入修订。最常用的版本是:连续 5 次读出 15 个名词(词表 A)要求受试者自由回忆,每次呈现的单词次序是固定的。每次读出单词之前重复指导用语,5 次后,读出干扰 15 个词语(词表 B),即刻自由回忆词表 B,紧接着回忆词表 A(第六次回忆),20 min 的时间间隔后,第七次回忆词表 A。再认有两种,一种是让受试者阅读一个故事,从中挑出词表 A 中呈现的单词,另一种是在词表 A、B 和 20 个语义或语音相似

的词语组成50个单词的词表中识别词表A的单词。

RAVLT评分记录纸(表4-1-1)的第一列英文单词是原来版本的记忆词表,在翻译时,我们考虑到受试者容易理解的因素,全部译为双字词语,比如,Drum,可以翻译为鼓,但是,受试者乍听之下,可能听成"谷"、"骨"、"蛊"等同音字,所以,我们翻译为"锣鼓"。在RAVLT中,不需要词语归类记忆,只要发音相同就算正确,所以,受试者听成另一个字并不影响得分;而在CVLT中,由于词语可以归类编码记忆,受试者正确理解词义是必要的。

表4-1-1 RAVLT评分记录纸

词表A		A1	A2	A3	A4	A5	词表B	B1	A6	A7	回忆词表A
Drum	锣鼓						Desk				Drum
Curtain	窗帘						Ranger				Curtain
Bell	时钟						Bird				Bell
Coffee	咖啡						Shoe				Coffee
School	学校						Stove				School
Parent	父亲						Mountain				Parent
Moon	月亮						Glasses				Moon
Garden	公园						Towel				Garden
Hat	帽子						Cloud				Hat
Farmer	农夫						Boat				Farmer
Nose	鼻子						Lamb				Nose
Turkey	火鸡						Gun				Turkey
Color	颜色						Pencil				Color
House	房子						Church				House
River	河流						Fist				River

注:来源 Strauss E, Sherman E, Spreen O. A compendium of neuropsychological tests (Third edition). Oxford university press. 2006, 776-780.

一、指导语

第一次学习时检查者说:"我要给你读一组单词,你仔细听,我读完后,你根据记忆尽可能多地复述这些单词,可以不考虑你复述单词的次序,现在开始,你回忆得越多越好。"然后,检查者读出词表A的15个单词,每个单词间隔1 s。核对回忆的单词,使用数字跟踪记录受试者的回忆单词的次序。对于受试者的回答是重复、错误数还是纠错数,检查者都不要给予反馈。在受试者表示他(她)再也回忆不出更多单词了,检查者说第二次指导语:"现在,我再读一

遍同样的单词，与刚才一样，我读完了要你回忆，回忆包括你刚才说过的词语。不用管你说的词语的次序。说得越多越好。”

第三至第五次的指导语与第二次相同。检查者可以表扬受试者回忆的词语越来越多，还可以告诉受试者每次回忆正确的词语有多少，这可以激发受试者的好胜心与自信心。第五次回忆之后，检查者读出词表B，指导语与词表A的第一次相同：“我要给你读第二组单词，你仔细听，我读完后，你根据记忆尽可能多地复述这些单词，可以不考虑你复述单词的次序，现在开始，你回忆得越多越好。”

紧接着回忆词表B，检查者请求受试者回忆词表A的词语，不要再读这些单词，“现在请你尽可能多回忆第一组词语。”在完成其他心理测验20 min的时间间隔后，检查者请求受试者回忆词表A的词语，不要再读这些单词，“现在请你尽可能多地回忆第一组词语。”

完成第七次回忆后，做再认测验。后者已见上述。如果要做词语次序判断，可以给受试者呈现一张15个词语随机排列的词语A，请求受试者写下原来听到的词语的次序号。

二、评分

对5次学习的总分、学习效率、前摄与倒摄抑制、保存与遗忘等指标评分。前摄抑制指第一次回忆对词表B回忆的影响；倒摄抑制指第五次回忆对第六次回忆的影响。其他指标包括：首因(primacy)效应：词表A的开头3个词语正确回忆数；近因(recency)效应：词表A的结尾3个词语正确回忆数；获得进入(gained access)指从一次回忆到下一次回忆的获得的单词，失去进入(lost access)指从一次回忆到下一次回忆的失去的单词，这2个指标可以判断历次回忆之间单词的获得与巩固情况。另外，Vakil与Blachstein(1994)提出一个附加项目，在没有预警的情况下，在测验结束时要求受试者写下词表A的单词呈现的次序。

根据不同年龄制定的RAVLT的美国人常模见表4-1-2。

表4-1-2　根据不同年龄制定的RAVLT的美国人常模(括号内为标准差)(Schmidt，1996)

年龄(岁)	*n*	学习*1*	学习*2*	学习*3*	学习*4*	学习*5*
50～59	161	6.2(1.6)	9.0(1.9)	10.5(1.9)	11.4(1.9)	12.1(2.1)
60～69	166	5.9(1.6)	8.4(2.0)	9.8(2.3)	10.9(2.3)	11.3(2.3)
70～79	143	5.5(1.6)	7.7(2.1)	8.8(2.1)	9.8(2.4)	10.3(2.4)
80～89	50	5.2(1.5)	7.2(1.8)	8.6(2.3)	9.7(2.3)	10.0(2.3)

（续表）

年龄(岁)	学习 *B*	学习 *6*	学习 *7*	再认	学习 *1～5*
50～59	5.7(2.2)	9.9(2.8)	9.9(3.2)	13.9(1.4)	47.6(8.1)
60～69	5.1(1.3)	9.3(2.9)	8.8(3.0)	13.5(1.3)	43.4(7.7)
70～79	3.9(1.6)	8.1(3.0)	7.0(2.4)	13.3(1.5)	37.1(7.5)
80～89	—	7.7(3.4)	—	13.0(2.3)	—

资料来源：Strauss E，Sherman E，Spreen O. A compendium of neuropsychological tests（Third edition）. Oxford university press. 2006,786－787.

三、评价

RAVLT 使用历史悠久，已经用于各种病因的神经科疾病的记忆功能评估，也用于糖尿病、慢性疲劳综合征、儿童学习障碍、重金属中毒、精神分裂症等的评估。不同的研究者使用 RAVLT 会有不同的形式，在词表长度、学习次数、间隔时间、是否采用词表 B、选择何种再认方式等方面各不相同。比如，有的研究者为了避免第 5 次学习后回忆呈现天花板效应而采用 3 次学习。

由于 RAVLT 的词语并没有语义关联，词语的次序标签是更重要的记忆策略，所以，RAVLT 比 CVLT 稍难。在检测记忆方面，CVLT 比 RAVLT 更敏感，也许与 CVLT 需要语义串联加工有关。比如，Stallings(1995)针对脑外伤患者的评估发现，CVLT 对于记忆障碍的检出率更高，与患者的记忆主诉的相关性也更高。

第二节　加利福尼亚词语学习测验(CVLT)

1989 年 Delis 编制了加利福尼亚词语学习测验(California verbal learning test，CVLT)第一版(CVLT－Ⅰ)，2000 年编制了 CVLT 第二版(CVLT－Ⅱ)。Delis 注意到，以词语为材料的记忆测验版本很多，但是，很少对提供的词语材料的加工处理过程与提取机制进行分析。CVLT 第一版(CVLT－Ⅰ)正是加工取向的研究(process-oriented approach)，因为可以提供不同类型的学习与记忆损害的剖面图，CVLT 得到广泛的应用。第二版与第一版不同的地方是词语的难度降低便于受试者理解、更大的正常人群数据库，分析方法也有更新。

一、CVLT-Ⅱ的操作步骤

针对词表A,检查者每秒1个词语读出单词,要求受试者连续5次学习,每次学习后要求受试者回忆。词表A由16个单词组成,可以分为4个语义类别(如家具类、蔬菜类、交通工具类、动物类),每个类别4个词语。呈现的词语是随机的,不按照语义类别排列。用于干扰的词表B也是16个词语,2个与词表A一样的类别,如蔬菜与动物类,2个不一样的类别,如乐器、房屋部件名称。词表B只学习一次,在词表A短延迟自由回忆之后。在非语言测验(如连线测验等)20 min后,要求针对词表A进行长延迟自由回忆、线索回忆和"是、不是"再认测验,大约10 min后,要求完成迫选再认测验。

CVLT-Ⅱ有一个简短版,叫CVLT-SF(short form),可以用于比较严重的患者或者作为初步筛查用。它包括9个单词,取消词表B,减少学习次数,缩短延迟间隔时间等。

二、CVLT-Ⅱ评分

CVLT-Ⅱ的评分有3套:核心报告(core report)、延展报告(expanded report)和研究报告(research report)。核心报告包括27个最常用的指标,延展报告有66个参数指标(CVLT-SF版是51个),对词语的学习与记忆功能进行深度分析。研究报告有260个参数指标,当然,大部分指标没有常模数据(表4-2-1)。

表4-2-1 CVLT-Ⅱ评分的主要变量及其定义

变量(汉语)	变量(英语)	定 义
List A total	词表A的总分	5次学习后回忆的词语正确数之和
List A1	词表A第一次回忆	第1次回忆的词语正确数
List A5	词表A第五次回忆	第五次回忆的词语正确数
List B	词表B回忆	干扰词表(List B)回忆的词语正确数
List A short delay free recall	词表A短延迟自由回忆	在干扰词表B回忆后立即回忆词表A的正确数
List A short delay cued recall	词表A短延迟线索回忆	在干扰词表B回忆后立即回忆词表A,给予语义线索后回忆的正确数
List A long delay free recall	词表A长延迟自由回忆	20 min延迟后自由回忆词表A的正确数
List A long delay cued recall	词表A长延迟线索回忆	20 min后、给予语义线索后回忆词表A的正确数

（续表）

变量(汉语)	变量(英语)	定　　义
Semantic clustering	语义串联	同一语义范畴连续回忆的个数，反映受试者利用语义组织词语的能力
Serial clustering	次序串联	根据词语呈现的一系列次序进行回忆
Primacy %	首因	词表 A 的开头部分词语正确回忆占总数的百分比
Middle %	中间	词表 A 的中间部分词语正确回忆占总数的百分比
Recency %	近因	词表 A 的结尾部分词语正确回忆占总数的百分比
Learning slope	学习速率	从每次学习中回忆的新的词语的平均数
Consistency %	一致性	历次回忆中相同词语数
Perseverations	持续性	在一次回忆中说出相同的正确词语的重复数
Free intrusions	自由回忆插入数	在历次自由回忆中，词表外的单词插入的个数
Cued intrusions	线索回忆插入数	在两次线索回忆中，词表外的单词插入的个数
Recognitions hits	再认击中	在是-不是形式再认测验中，属于词表 A 的词语数
Discriminability	区分力	再认测验中区分目标词语与分心的干扰词语的准确性
False positives	假阳性	再认测验中未能正确识别词表 A 的目标单词的个数
Response bias	反应偏差	再认测验中反应风格，是倾向于把错误的说成对的，还是倾向于把对的说成错的

这些参数指标可以分为 14 大类：① 历次回忆的粗分和总分；② 学习策略，如语义串联、主观组织；③ 首因（primacy）效应和近因（recency）效应；④ 每次回忆表现的学习速率；⑤ 历次回忆中相同词语反映的一致性；⑥ 前摄与倒摄干扰程度；⑦ 经过延迟间隔后词语的保存情况；⑧ 类别线索与再认对回忆的提高幅度；⑨ 根据信号识别理论再认的破坏情况；⑩ 词语编码、储存与提取过程的完整性；⑪ 回忆的插入错误类型，如语义相关的、语义无关的；⑫ 回忆的重复错误；⑬ 再认的假阳性错误类型分析；⑭ 完成测验的心理努力程度。

三、评价

大量研究证实 CVLT 识别记忆损害的敏感性不仅优于 RAVLT 和选择提醒测验(selective reminding test, SRT),也优于韦氏记忆测验修订版(WMS-R)。通过 CVLT 检测头部外伤、癫痫、AD、帕金森病、亨廷顿病、缺血性血管性痴呆、柯萨可夫综合征、艾滋病、抑郁症和精神分裂症等不同疾病,可以发现特征性的记忆和学习损害的剖面图,从而有效区别不同疾病所致认知功能减退。如有无左侧海马硬化的受试者在首因和近因效应方面有显著差异;亨廷顿病患者的记忆保持率较高但词语重复较多;根据使用的编码策略的差异可以将头部外伤患者的记忆缺损区分为 4 种类型等。

第三节　听觉词语学习测验(AVLT)

听觉词语学习测验(auditory verbal learning test, AVLT)的国内版本目前有 4 种,主要是词语材料不同,操作过程和得分分析方法相似。① 上海精神卫生中心版本(改编自 WHO-AVLT)采用 15 个词语:手臂、耳朵、眼睛、猫、狗、马、斧、刀、锤子、床、闹钟、椅子、飞机、自行车、汽车。② 香港大学-安徽医科大学版本(HKU-AHMU,改编自 RAVLT)采用的 15 个词语是:鼓、窗帘、门铃、咖啡、学校、父亲、月亮、公园、帽子、农民、鼻子、母鸡、颜色、房屋、河流。③ 香港中文大学版本(HKVLT,改编自 CVLT)16 个词语是:祖母、伯父、表弟、侄女、镜子、书桌、衣柜、电灯、印度、智利、泰国、瑞士、番茄、黄瓜、花菜、洋葱。④ 华山医院版本(华山版,AVLT-H,1998 年参考 CVLT 与 HKVLT 编制)采用 12 个词语:大衣、长裤、头巾、手套、司机、木工、士兵、律师、海棠、百合、蜡梅、玉兰,此外,有 3 处修改:词语重复学习改为 3 次;3 次学习后告知受试者记住这些词语、后面还要回忆这些词语;删除 16 个干扰词语回忆;增加间隔 3～5 min的"短延迟回忆",即在非言词测验间隔 3～5 min 后,回忆刚才的 12 个词语。因为原版本用于痴呆患者,其延迟自由回忆得分极低,经常是 0 分,呈现地板效应,增加"短延迟回忆"有助于观察受试者得分衰减过程。华山医院版本与 2001 年出版的霍普金斯词语学习测验(HVLT-R)在词语数量和操作步骤方面基本相同。

一、AVLT－H 的编制

AVLT 汉语版本词语的选择(RAVLT)和 SRT 的词语选择要求在意义上互不相干，CVLT 则相反，要求选择的词语使受试者尽可能按照语义归类记忆，借此分析受试者记忆过程中的内在编码情况。因为对于痴呆的早期识别，语义归类记忆比其他归类编码方式更敏感，所以，在词语的选择中，应根据汉语词语的特点，注意是具体名词还是抽象名词、汉字的字数(形成记忆的字数编码)、词性(形成词性编码)、有无同音字(形成语音编码)、汉字笔画数(通过视像化形成字形编码)、熟悉性(冷僻词或假词会使受试者听不懂而不能形成语义编码，过于熟悉，如眼睛、鼻子、耳朵、嘴巴等则容易形成内隐编码)。我们选用的词语，包括服装类、职业类和花朵类名词，每类 4 个名词共 12 个词汇随机组成。再认词语中干扰词的选择包括同类、读音近似者、同类兼读音近似者、和意义与读音均无关者。

二、AVLT－H 的评分

AVLT－H 主要变量包括每次回忆正确数(包括即刻回忆、"短延迟回忆"、"长延迟回忆"、线索回忆和再认)和错误数；学习能力；记忆保持率；辩正能力；概念记忆：又称为类别记忆、语义串联记忆(深加工记忆，反映语义编码程度，连续 2 个同类名词作为语义串联 1 分，连续 3 个同类名词作为语义串联 2 分，全部按照语义串联回忆，得 12 分)；主观组织(浅加工记忆)；首因和近因效应(primacy and recency effects：首因效应指每次回忆中前 4 个词回忆的数目，近因效应指每次回忆中后 4 个词回忆的数目)；反应偏差(response bias，即在再认测验中，是倾向于把错误的说成对的，还是倾向于把对的说成错的)。不同年龄组中国上海城区中老年人的 AVLT－H 表现见表 4－3－1。

表 4－3－1　不同年龄组中国上海城区中老年人的 AVLT－H 表现

指　标	50～59 岁 (*n*=420)	60～69 岁 (*n*=376)	70～79 岁 (*n*=316)	80～89 岁 (*n*=103)
第一次学习	4.00(1.62)	4.06(1.85)	3.62(1.63)	3.06(1.47)
第二次学习	6.25(1.93)	6.17(1.87)	5.62(1.75)	5.00(1.39)
第三次学习	7.61(2.00)	7.43(2.05)	6.93(1.78)	6.37(1.76)
短延迟回忆	6.57(2.05)	6.16(2.17)	5.56(1.97)	5.01(1.77)

（续表）

指　　标	50～59岁 (n=420)	60～69岁 (n=376)	70～79岁 (n=316)	80～89岁 (n=103)
长延迟回忆	6.44(1.89)	6.00(2.19)	5.24(1.96)	4.80(1.91)
线索回忆	6.96(3.24)	7.03(3.11)	6.33(2.95)	5.62(3.12)
再认	11.10(1.2)	10.80(1.4)	10.70(1.4)	19.50(3.27)
前3次学习之和	17.87(4.69)	17.69(4.83)	16.18(4.42)	14.45(3.70)

资料来源：郭起浩，孙一忞，虞培敏，洪震，吕传真. 听觉词语学习测验的社区老人常模. 中国临床心理学杂志，2007，15(2)：132－135. 并结合2007～2012年数据重新统计结果。

三、AVLT－H的评价

与CVLT、ReyVLT相比，AVLT－H的词语数量从15～16个减少到12个、学习次数从5次减少到3次，难度似乎有所降低，但是，已经有比较研究表明敏感性并没有降低。

目前AVLT最重要的应用是识别轻度认知损害(MCI)。MCI的识别对于AD的早期诊断、早期治疗有重要意义。MCI研究用诊断标准要求“有记忆减退的客观证据”，通常用AVLT的延迟回忆得分少于年龄和教育程度匹配组的“均数－1.5标准差”来表示。延迟回忆被认为是AD认知功能损害最早、最敏感的指标。Tierney通过对123名有记忆损害主诉的非痴呆老人随访2年，有24%发展为AD，分析基线时样本的神经心理测验表现，以CVLT的延迟记忆得分最有意义，预测准确性为89%。Visser编制临床前AD诊断量表(PAS)由患者年龄、MMSE总分、总体严重度量表、认知测验、影像学呈现颞叶内侧萎缩和APOE基因型6个部分组成，其认知测验由听觉词语学习测验和1～3种其他认知领域测验(如Stroop色词测验)组成。CVLT的词语延迟回忆在认知下降(MCI转化为AD)和认知稳定(MCI未转化为AD)两组间最具鉴别力已经被多个纵向调查所证实。

我们应用AVLT－H检查正常老人，发现高教育的老人的延迟回忆优于短时回忆，即随着时间间隔延长，记忆成绩不是下降而是提高，这种反跳现象(rebound phenomenon)在AD患者和遗忘型MCI患者中是没有的。教育程度低或文盲老人的AVLT表现差异比较大，标准差甚至大于均数，不能用于低教育老人MCI的识别是AVLT的主要缺点，为了弥补这一缺点，我们还编制了AVLT的图片版，即以图片显示记忆材料，测验过程与分析策略和AVLT相似，图片短时记忆和延迟回忆识别MCI也相当敏感，且测验员之间

一致性和可接受性(完成率)更佳,但是目前国际上使用 AVLT 图片版的还非常少。

第四节　霍普金斯词语学习测验修订版(HVLT-R)

霍普金斯词语学习测验修订版(Hopkins verbal learning test-revised, HVLT-R)是 Brandt 和 Benedict(2001)编制的,在检测方法上与 Benedict 1997 年出版的 BVMT-R(Brief visuospatial memory test-revised)相似,即使中度痴呆患者依然可以使用。12 个名词的词表,3 个语义类别(如 4 腿动物类、宝石类、服装类),每个类别 4 个单词。HVLT-R 由 3 次学习、延迟回忆(在没有预先提示需要回忆的情况下间隔 20~25 min 后回忆)和再认,再认包括 12 个学习过的目标词与 12 个没有学习过的干扰词,干扰词中 6 个与目标词是同一类别的。很显然,HVLT-R 与 CVLT、RAVLT 在结构上是相似的,只是HVLT-R 更简短。2002 年,Benedict 发表的资料中增加了再认前的线索回忆环节。

一、评分方法

3 次学习和延迟回忆的满分分别是 12 分,总分是第一次到第三次的回忆之和,保存百分比(% Retained percentile)是第四次回忆除以第二和第三次回忆(×100%),再认分辨指数(discrimination index)是真阳性减去假阳性。进一步细致分析,可以将再认区分为"真阳性个数、语义相关假阳性错误个数、语义无关假阳性错误个数、假阳性错误总数"。澳大利亚老人的 HVLT-R 常模见表4-4-1;非裔美国人的 HVLT-R 常模见表 4-4-2。

表 4-4-1　澳大利亚老人的 HVLT-R 常模 (Hester,2004)

分析指标	60~69 岁		70~79 岁		80~89 岁	
	教育≤10 (n=29)	教育≥11 (n=35)	教育≤10 (n=63)	教育≥11 (n=45)	教育≤10 (n=15)	教育≥11 (n=16)
第一次学习	5.2(1.5)	6.4(1.7)	4.7(1.9)	5.3(1.5)	4.2(1.5)	5.1(1.3)
第二次学习	6.7(2.2)	8.5(2.0)	6.8(2.0)	6.7(1.9)	6.1(1.8)	7.4(1.6)
第三次学习	8.1(2.3)	9.8(1.6)	7.8(2.5)	8.2(2.0)	7.1(2.5)	8.6(2.1)
学习能力	2.9(1.8)	3.5(1.5)	3.2(1.8)	3.1(1.8)	3.2(1.6)	3.8(1.3)

(续表)

分析指标	60～69岁		70～79岁		80～89岁	
	教育≤10 (*n*=29)	教育≥11 (*n*=35)	教育≤10 (*n*=63)	教育≥11 (*n*=45)	教育≤10 (*n*=15)	教育≥11 (*n*=16)
前3次学习之和	20.0(5.5)	24.6(4.8)	19.4(5.8)	20.2(4.6)	17.4(5.2)	21.1(4.6)
延迟回忆	6.3(3.3)	8.4(2.9)	6.4(3.5)	7.3(2.7)	5.4(3.1)	5.4(2.6)
保存百分比	73.7(34.8)	82.9(23.3)	80.4(36.3)	83.6(23.9)	78.3(48.8)	57.9(23.6)
再认分辨指数	8.4(2.3)	9.9(1.8)	8.9(2.2)	9.4(2.0)	9.0(2.3)	8.6(3.1)

资料来源：Strauss E, Sherman E, Spreen O. A compendium of neuropsychological tests (Third edition). Oxford university press. 2006,762.

表4-4-2 非裔美国人的HVLT-R常模(Friedman,2002)

分析指标	60～71岁(*n*=111)		72～84岁(*n*=126)	
	均值(SD)	范围	均值(SD)	范围
第一次学习	4.4(1.3)	2～8	3.8(1.5)	0～8
第二次学习	6.3(1.3)	4～11	5.4(1.5)	2～9
第三次学习	7.2(1.4)	4～11	6.3(1.7)	3～11
学习能力	2.9(1.1)	0～6	1.2(2.6)	2～5
前3次学习之和	17.9(3.5)	11～30	15.5(4.3)	7～28
延迟回忆	6.6(1.6)	3～11	5.8(2.0)	1～10
线索回忆	7.2(1.6)	3～12	6.4(2.0)	2～11
真阳性	11.1(1.3)	7～12	10.1(2.1)	4～12
假阳性	1.6(1.1)	0～4	1.8(1.3)	0～6
区分力	9.4(1.9)	3～12	8.4(2.7)	0～12
保存百分比	90.5(15.0)	50～140	88.8(17.9)	17～133

资料来源：Strauss E, Sherman E, Spreen O. A compendium of neuropsychological tests (Third edition). Oxford university press. 2006,764.

二、评价

HVLT-R的基本特征：① 年龄：年龄影响HVLT-R的变量，可以解释总分的19%，保存的3%。教育可以解释总分的5%，其他指标的影响比较小。女性的总分超男性3%。对于年轻或中年的受过良好教育的受试者，HVLT-R有可能出现天花板效应，尤其是第三次回忆。② 信度：Benedict(1998)报道40例老人间隔6周复测，相关系数：总分0.74；延迟回忆0.66；保存百分比0.39；再认分辨指数0.40。③ 练习效应：2周后复测，总分提高约3分。但是，学习曲线是相似的。④ 效度：聚合效度与CVLT相似。学习曲线与保存

率非常接近，与痴呆严重度的相关性也接近。

HVLT－R用于痴呆识别是有效的，如 Shapiro(1999)报道 HVLT－R 总分低于 1SD 时作为分界值，识别痴呆的敏感性 95%，特异性 83%。Hogervorst (2002)综合 HVLT－R 的总分与再认分辨指数，识别痴呆的敏感性优于 MMSE。HVLT－R 用于区分 AD 与 VaD 的效果不佳，但可以用来区分 AD 与亨廷顿病，AD 患者的再认有损害，而亨廷顿病患者尽管自由回忆受损，再认是完整的。由于 HVLT－R 有 6 个难度相当的版本，在需要系列评估的临床研究，比如，在与头外伤、情绪障碍、营养缺乏症和药物临床试验相关的认知功能评估方面已经得到应用。Pardee (2005)发现受试者在他不喜欢的时间段完成的HVLT－R比在他喜欢的时间段完成略差。

总之，HVLT－R 用于识别痴呆是简便有效的，但是，对于非老年的、轻微或(和)复杂的认知损害建议采用CVLT－Ⅱ。

第五节　选择提醒测验(SRT)

布谢克选择提醒测验又称为 SRT 或言语选择提醒测验(verbal selective reminding test，VSRT)，是 1973 年 Buschke 首先编制的，后来的版本非常多，因为每个研究者自己编制了一套词语。SRT 由 12 个意义不相关的词语组成，每次呈现后回忆 1 次，检查者提醒有哪些词语没有回忆，下一次也一样提醒没有回忆的词语，一共可以学习 12 次，或者连续 3 次回忆全对。学习后的线索回忆是提供 2～3 个起首字母要求说出词表的相应的单词。接着线索回忆的是多选再认，检查者连续呈现 12 张卡片，每张包括一个词表中的词语、一个同义词、一个同音异义词和一个无关干扰词语，最后，是 30 min 间隔的延迟回忆，事先不要告诉受试者需要回忆。有各种词语版本，表 4－5－1 是其中的一种。

表 4－5－1　SRT 评分记录纸

目标词语	1	2	3	4	5	6	7	8	9	10	11	12	CR	MC	30 min
Bowl															
Passion															
Dawn															
Judgement															
Grant															

（续表）

目标词语	1	2	3	4	5	6	7	8	9	10	11	12	CR	MC	30 min
Bee															
Plane															
County															
Choice															
Seed															
Wool															
Meal															

资料来源：Strauss E, Sherman E, Spreen O. A compendium of neuropsychological tests (Third edition). Oxford university press. 2006,714.

一、评分方法

在编制中文版本的再认测验部分，显然不能直接引用英文版的翻译（表4－5－2），因为同音异义词翻译后就没有同样的干扰作用了。其实，英文版中，聪明的受试者从这4个词语的安排中有一定的概率能猜到正确答案。SRT的评分方法见表4－5－3。

表4－5－2 SRT的目标词语与再认词语及其翻译（使用时采用随机次序）

目标词语		同义词		同音异义词		无关词语	
Bowl	碗	Dish	碟子	Bell	钟	View	视图
Passion	热情	Love	爱情	Poison	毒物	Conform	包裹
Dawn	黎明	Sunrise	日出	Down	下降	Bet	打赌
Judgement	判断	Verdict	裁定	Fudge	捏造	Pasteboard	硬纸板
Grant	同意	Give	给予	Grand	伟大	Jazz	爵士
Bee	蜜蜂	Sting	叮咬	See	看见	Fold	重复
Plane	飞机	Jet	喷气式飞机	Pain	疼痛	Pulled	行驶
County	郡	State	州	Counter	柜台	Tasted	体验
Choice	选择	Select	挑选	Cheese	乳酪	Voice	声音
Seed	种子	Flower	花朵	Seek	寻找	Herd	畜群
Wool	羊毛	Sheep	绵羊	Would	愿意	Date	日子
Meal	膳食	Food	食物	Mill	工厂	Queen	王后

表 4-5-3 SRT 的评分方法

简称	意义	定义
LTS	长时储存	假如一个词语连续 2 次回忆上都呈现，可以假定这个单词已经进入长时储存
LTR	长时提取	受试者回忆已经进入长时储存的词语，可以作为长时提取
CLTR	持续长时提取	当词语在历次回忆中均能呈现，可以称为持续长时提取
RLTR	随机长时提取	不一致的长时提取，长时储存中的单词在后面的回忆中忘记了，是随机长时提取
STR	短时回忆	没有进入长时储存的词语个数
Sum recall	总分	每次回忆的 STR 与 LTR 之和
Reminders	提醒数	12 次学习的总提醒数，总分=144

有许多研究者将上述版本的 12 个词语的学习次数减少，比如，8 次或连续满分 2 次，用于青少年受试者。关于人口学效应，大部分 SRT 指标与增龄相关，女性优于男性。关于教育程度的影响报道不一致。

二、评价

Sass(1990)报道 SRT 得分与左侧(而不是右侧)海马锥体细胞密度显著相关，然而，也有报道左额叶，甚至右半球病灶引起 SRT 的得分下降，所以，不能根据 SRT 的得分预测海马的异常程度。与 RAVLT、CVLT 一样，SRT 的延迟回忆是 30 min，Bell(2005)报道一组颞叶癫痫患者采用 SRT 测验时分别要求 30 min 与 24 h 延迟回忆，结果发现患者组与正常对照组的在更长的延迟后记忆下降幅度是相似的，超过 24 h 的遗忘并没有加速。LTR 与 CLTR 在识别轻度痴呆方面比 SRT 的其他指标更有价值。SRT 也可以用于预测痴呆。SRT 的一个优点是在神经科疾病的检测中没有学习效应，即使再次检测仅仅间隔几天时间，但在正常受试者中有学习效应。对于比较重的患者，可以采用 6 个词语的删减版本。

第六节 逻辑记忆测验(LMT)

逻辑记忆测验(logical memory test，LMT)是韦氏记忆测验(Wechsler memory scale, WMS)的一个分测验，是对一段或一串小故事的回忆。近 10

年，因为在MCI的诊断标准中它被认为是客观记忆受损的指标之一而得到广泛的应用。

一、概况

与词语学习测验相比，LMT具有以下特点：LMT更接近现实生活，平时看到的电影电视、文学作品及日常交谈，无一不是故事；故事更容易组织编码，更容易记忆；词语学习测验呈现的词语间隔时间是固定的，而阅读故事的速度不容易把握，施测者阅读故事的速度直接影响受试者的回忆成绩。

LMT作为韦氏记忆测验的一个分测验，在其初版中只有即刻回忆，1987年的修订版有一个重要的变化，是增加了30 min后的延迟回忆。作者还开发了有12个问题的线索回忆版本和多选回忆版本。多选回忆的再认任务对于高教育受试者容易出现天花板效应。

二、指导语

“我读给你听一段小故事，只有短短的几行，你仔细听，尽可能记住它，并且要尽可能记同一词语。我读完后，我要求你告诉我刚才听到的内容。即使你不肯定的内容，你也可以告诉我。准备好了吗?”

值得注意的是，英文版本的LMT是听觉呈现的，而中文修订版(WMS-RC)中的故事是视觉呈现的，不同的感觉呈现方式对于识别MCI的敏感性有明显差别。作者主张国内在使用LMT时也采用听觉呈现方式。

三、评分

因为很少有受试者精确地复述故事材料，评分方法存在许多问题。常见的错误包括替换(同义词、相似概念、数字不准等)、遗漏、添加和细化、故事次序的转移等。

回忆故事的风格有两种：逐字的、一字不差的回忆；观念的、“得意忘形”的回忆，有些患者回忆了大量的细节，由于不是一字不差的回忆而得不到分数，还有些患者对故事本身已经没有连贯的记忆，但是能够一字不差地回忆许多细节而获得高分。Talland等认为可以通过受试者的同义词替代或适当的词组与精确单词的比例来区分这两种回忆风格。

为了避免评定员打分标准的不一致，WMS制定了明确的详细的评分标准：如“警察”说成“武警”、“咖啡店”说成“茶叶店”等如何评分均有指导手册

规定。近来针对主题而非细节的评分标准也已推出。

四、评价

LMT 历史悠久、普及，所以，几乎所有脑器质性疾病都使用过 LMT 的评估。Martin(1999)发现未手术的癫痫患者左海马容积可以预测 LMT 的即刻回忆、延迟回忆和储存百分比的得分。Webster(1992)采用"重要"、"细节"和"插入反应"3 个评分的比例来判断病灶的一侧性，右半球病变有更多的插入错误，而左半球病变在这 3 个指标上都比较差。LMT 也可以用于识别多发性硬化、闭合性脑外伤等患者的记忆受损。

近 10 年，LMT 最常用于 MCI 的诊断，比如 ADNI 对 3 组研究对象的入组标准(表 4-6-1)中，LMT 延迟回忆(LMT-Ⅱ)得分是其核心指标。

表 4-6-1　ADNI 3 个研究对象的入组标准要点

指　　标	正常老人	遗忘型 MCI	轻度 AD
MMSE	24～30 分	24～30 分	20～26 分
CDR	0 分	0.5 分	0.5 分或 1 分
LMT-Ⅱ：教育16 年以上	9 分以上	8 分以下	8 分以下
8～15 年	5 分以上	4 分以下	4 分以下
0～7 年	3 分以上	2 分以下	2 分以下

附件 4-1　英文故事

英文故事 1

住在南 Boston 的 Anna/Thompson，受雇在学校咖啡店做厨师，报告夜晚在 City Hall 车站，State Street 被抢劫了 56 美元。她有 4 个小孩、要付房租，已经 2 天没有吃的了。警察很同情这位女性的遭遇，纷纷为她捐款。

中文故事 2(括号内为评分要点)

从前有一个[国王]，要他的[大臣]在[明天]上朝时献上[公鸡蛋]。大臣很[着急]，因为知道[没有]公鸡蛋。大臣有一个[儿子]，[12 岁]，[知道]了这件事，[安慰]了父亲，[自己]去见国王。对国王说[父亲]在家正要[生孩子]了，所以[不能]上朝。国王很[生气]，说[男人]哪能生孩子？儿子说你[既然知道]男人不能生孩子，那么，[为什么]要公鸡生蛋？国王自知[理屈]，便[不再]提公鸡蛋。

中文故事 3(括号内为评分要点)

从前有一[青年]，有一次出[远门]，[过河]时坐在[船边]，将[宝剑][落入

河里]，[船夫]为他[惋惜]，[停住船]，[叫]他[下水][打捞]。青年说，[不用急]，我已在[船沿]做了[记号]，[等到]船[靠岸]，[我从有]记号的地方[下水]，[准可]捞到。

（郭起浩）

第七节 Rivermead行为记忆测验(RBMT)

Rivermead行为记忆测验(Rivermead behavioural memory test，RBMT)评估患者的日常记忆功能受损，并监测患者记忆障碍的变化。RBMT-Ⅱ标准版包括4个平行测验、1本手册、2本附录、25份记分表、1盒录音带及1个闹钟。另外还有难度更高的扩展版(RBMT-E)和难度略低的儿童版(RBMT-C)。已经有荷兰语、德语、西班牙语、中文及日文等语言版本。

一、RBMT版本介绍

(一) RBMT/RBMT-Ⅱ

RBMT是1985年Wilson等为获得脑损伤患者记忆障碍的康复干预评估工具而设计的。评定项目没有遵守特别的记忆理论模型，而是试图模拟日常生活中需求的记忆。评定项目包括日常生活中常见的记忆问题，也是生活所必须记住的信息类型。2003年发表的新版(RBMT-Ⅱ)针对第一版做了一些细微的改变：① 面孔照片系列增加了不同种族的面孔以适应多种族社会；② 路线记忆的评分略有变化；③ 包装由纸质改为塑料包装。

评定内容包括11个项目，选自脑损伤患者记忆问题的研究报告，包括记姓和名、记所藏物品、记约定、图片再认、故事回忆(即时和延迟)、面孔再认、路线回忆(即时和延迟)、信件回忆(即时和延迟)、定向和日期(表4-7-1)。

表4-7-1 RBMT的内容

项目编号	中文名称	英文简称	内容
1,2	记姓和名	First and Second Names	呈现一肖像照片，要求20分钟后回忆其姓名
3	记所藏物品	Belonging	被试的物品，测验结束后回忆藏在哪里
4	记约定	Appointment	间隔20分钟闹钟响的时候，回忆约定的句子
5	图片再认	Picture Recognition	回忆20分钟前呈现的20幅线条图片中的10幅
6	故事回忆	Story	即刻回忆与20分钟延迟回忆21个要点的故事

（续表）

项目编号	中文名称	英文简称	内　　容
7	面孔再认	Face Recognition	从 10 张呈现的面孔中识别 5 张面孔
8	路线回忆	Route	即刻回忆与 20 分钟延迟回忆路线
9	信件回忆	Messages	即刻回忆与 20 分钟延迟回忆路线左边的信件
10,11	定向和日期	Orientation and Date	回答有关定向与日期的 10 个问题

情景记忆为常见物品及面孔的再认。前瞻记忆包括下面 3 个项目：① 记住测试前拿走的个人物品；② 警铃响的时候能回忆约定的问题句；③ 记住一段短的路线。定向包括时间、地点、人物定向。RBMT/RBMT－Ⅱ拥有 4 个不同的平行版本 A、B、C、D，用来避免学习效应。

（二）RBMT－C

儿童版的 RBMT 除故事回忆及定向问题不同以外，其他的与成人版相同。另外，第一、第二项（记住姓和名）及第四项（记约定）不适用于 8 岁以下儿童。

（三）RBMT－E

RBMT 是一项筛查量表，对轻度损害诊断的敏感度较低。为了增加测试对轻度损害诊断的敏感度，Wilson 等将测试中所需记忆项目的数量翻倍。原版中 4 个平行测验中的 A 版及 B 版合并为 RBMT－E 第一版，C 版及 D 版合并为 RBMT－E 第二版。其中故事回忆未改变，因为它在正常人中也没有明显的天花板效应。另外，为了避免天花板效应及地板效应，部分项目做了一些小的变化，如增加了额外的面孔记忆材料。

二、评估方法

针对语言或知觉能力有缺陷的受试者，手册中提供了 RBMT 的简略版本；针对语言能力缺陷的受试者，去除了记姓和名、定向及故事回忆；针对知觉能力有缺陷的受试者，去除了路线回忆、面孔再认、定向及日期。简略版本可能会过高地估计受试者的记忆损害。11 个项目分 16 个步骤合理安排、相互穿插评定，要求按量表规定的顺序进行。

RBMT 路线回忆要求受试者在医生示范后即时和一段时间后重复一段短路线。这对行动不便的受试者来说有一定难度。针对这部分受试者，Towle 等提出了替代的评定方法。受试者可在房间图中画出路线，或在模型（有树、房子、桥、车库）中移动玩具车。Clare 等认为 RBMT－E 路线回忆和信件回忆

可以改为只需模型操作的模型路线即刻与延迟回忆和报纸即刻与延迟回忆。

RBMT-Ⅱ评估耗时约25 min。评分标准如下。

(一) RBMT/RBMT-Ⅱ

每一项都由初步积分换算成筛选分数(screening score)(0-错误,1-正确)和标准分数(profile score)(0-异常,1-可疑,2-正常),之后计算总分。筛选分数和标准分数满分分别为12分和24分,其中筛选分数为粗量表分,提示受试者是否有行为记忆损害;标准分数为精量表分,提示受试者行为记忆损害严重度及具体损害类型。

RBMT-Ⅱ中路线即时回忆和延迟回忆的初步积分较RBMT有轻度差别,如筛选分数最高分由5分变为11分。Wilson等(2003)对比了两种回忆的评分方法发现两种不同方法中结果异常、可疑、正常的比例相似。

(二) RBMT-C

每一项都由筛选分数换算成标准分数(0-异常,1-可疑,2-正常),之后计算总分。5~6岁、7岁、8~10岁儿童的满分分别为18分、20分和22分。

(三) RBMT-E

每一项都由筛选分数换算成标准分数(0~4),之后计算总分,满分48分,分为4个等级(0-行为记忆能力受损,4-行为记忆能力完好)。其中某些项目的标准分受被试年龄及临床前IQ(由成人阅读测验NART、语言运用速度和能力测试SCOLP或韦氏成人阅读测试WTAR评定)的影响。

如RBMT-E采用模型路线回忆和报纸回忆,模型路线回忆积分共19分(9个路线点正确1个得1分,起始点正确得1分,结束点正确得1分,8个路线顺序正确1个得1分),报纸回忆积分共6分(3个项目,每个项目中,能自觉地拾起物品,可得2分;经提示拾起物品,可得1分;物品放置于正确地点,可得另外1分)。

三、人口统计学影响

1. 年龄　RBMT/RBMT-Ⅱ分数受年龄影响,约8岁可达到成人水平,中年后得分将会有所下降,也有研究认为分数下降只在70岁以后出现。在RBMT-E中,年龄影响受试者故事延迟回忆、路线即时和延迟回忆、信件延迟回忆、记约定和记所藏物品的表现。年轻者优于年老者。

2. IQ　智商对RBMT有轻到中度影响(r=0.21~0.58)。在RBMT/RBMT-Ⅱ中,智商主要影响定向及故事延迟回忆得分。在RBMT-C中,智

商主要对低年龄者（5～6 岁）有影响，而对高年龄者（7 岁以上）无影响。在 RBMT - E 中，智商对大部分项目均有影响，如故事回忆、图片再认、路线回忆、定向、记姓和名。

3. 教育程度　教育程度也影响 RBMT 表现。

4. 性别　大部分学者认为性别对测试无影响，仅个别研究认为男性记所藏物品和路线即时回忆的表现优于女性。而在 RBMT - E 中，女性的路线延迟回忆、信件即时和延迟回忆、记姓、记约定和记所藏物品的表现优于男性。

5. 种族　Wilson 等（1999）在 RBMT - E 中未发现欧裔、亚裔、加勒比-非裔人种之间有差异。

四、常模

最初 RBMT/RBMT - Ⅱ 常模是根据 176 位脑损伤患者（平均年龄 44.4 岁）和 118 位正常人（年龄 16～69 岁，平均 41.17 岁；IQ 68～136，平均 106）建立的。但这个常模的用处不大，大部分脑损伤组患者的得分都低于正常组 5 个百分点。所以笔者建议根据临床表现划分记忆障碍严重程度（正常、轻度、中度、重度）。筛选分低于 10 分、标准分低于 22 分被认为异常。低 RBMT/RBMT - Ⅱ 分数主要提示受试者智商受限或记忆障碍，在解释分数时需要注意受试者的智商。

Fraser 等（1999）测试了 131 位有独立生活能力的正常新西兰老人（年龄 60～89 岁，平均 72.71 岁；教育程度 6～18 年，平均 10.46 年；无记忆障碍主诉，无心血管及呼吸疾病史；MMSE 得分均高于标准值），建立了不同年龄组（60～69 岁、70～79 岁、80～89 岁）的常模。

青少年的表现与成人类似。笔者认为成人的常模可直接应用于智商正常及以上的青少年，而针对智商低于正常的青少年，筛选分数可降低 1 分，标准分降低 2 分。

RBMT - E 常模是根据 191 位英格兰人和澳大利亚人（年龄 16～76 岁，平均 39.4 岁，标准差 15.45）的测试结果建立的。但是样本量太小，年龄范围大，纳入和排除标准未介绍。笔者建议每项标准分 2～4 分（总分 24～48 分）提示记忆能力正常。但要注意的是，如果根据这个标准，25%的正常人将出现假阳性。

五、信度

1. 内部一致性　中文版 RBMT 具有较高的克隆巴赫系数（Cronbach's

alpha;0.86)(Man 和 Li,2001),其他版本的 RBMT 未检测。

2. *重测信度* Man 和 Li(2001)分别于基线及 2 周后对脑梗死患者进行了中文版 RBMT 测试,发现其具有较高的重测信度(r=0.89)。

3. *复本信度*

(1) RBMT/RBMT-Ⅱ:研究发现,四个平行测验版本之间相关系数范围为 0.67~0.88,只有一个大于 0.80;筛选分数之间的相关系数为 0.78,标准分数之间的相关系数为 0.85;毫无疑问,标准分更能反映受试者的能力。第二次测试分数,特别是记所藏物品,相关系数较第一次稍高。但四个版本间的难度是否有差别没有报道。Fraser 等(1999)对 26 位老年人进行了版本 A 测试,对 105 位成人进行了版本 B 测试,未发现两个版本间标准分数(相关系数 0.94)和每个项目的初步积分有差别。Man 和 Li(2001)报道了各版本之间初步积分的相关系数为 0.38~0.92。Wilson 等(2003b)对 85 位青少年进行了两个版本的测试(测试间隔时间未交代),未发现两个版本间的筛选分数和标准分数有差异。

(2) RBMT-C:Aldrich 和 Wilson(1991)对儿童进行了两个版本的测试(AB,BC,CD 或 DA),间隔时间为 48 小时,发现儿童版的相关系数与年龄相关,范围为 0.44~0.73。更低的相关系数可能反映了大年龄组的天花板效应。两次测试间未发现学习效应。

(3) RBMT-E:191 位正常人完成了两个版本的 RBMT-E,6 位只完成了其中的一个版本。197 位正常人平均年龄 39.4 岁(标准差 15.45)。大约一半人先完成第一版测试,至少 1 周后再完成第二版测试,另一半则相反。发现两个版本的故事即时回忆、路线即时和延迟回忆、记姓的得分有差异。这些差异(小于 1 初步积分)都需要在量化评分中考虑到。

4. *评估者间信度* RBMT 筛选分和标准分的评估者间信度为 100%(Wilson 等)。Man 和 Li(2001)报道了中文版 RBMT 评估者之间的信度为 0.74~0.95。

六、效度

1. *标准效度* RBMT 与记忆测验(沃灵顿再认记忆测验、韦氏记忆测验、Luria-Nebraska 记忆量表、Rey 图形回忆、Rey 听觉词语学习测验)具有中等相关性。RBMT 与注意测验(数字广度测验)相关不大,但在 7~9 岁儿童中可发现中等相关性。RBMT 与 MMSE 的相关系数为 0.85,与韦氏成人智力量表

修订版（WAIS-R）的相关系数为0.57。Malec等（1990）报道了RBMT与WAIS-R中言语理解因素及执行功能障碍无相关性，与知觉组织因素中等相关。这提示RBMT主要评估患者的记忆功能（新知识的学习和延迟回忆）。

2. 因子分析　Efklides等（2002）报道了RBMT、WMS及日常记忆问卷（EMQ）的验证性因子分析，包括言语记忆因子（RBMT故事回忆、WMS逻辑记忆）、学习因子（RBMT记姓名及约定、WMS联想学习）、定向因子（RBMT定向和日期、WMS个人信息及定向）。RBMT视觉再认项目不包含WMS视觉再生项目。纯粹的前瞻记忆成分（影响记所藏物品、记约定、信件回忆）在AD组及正常组均未确定。Maylor（1995）指出测验不是自发的，路线回忆和信封回忆项目中受试者需要在观察后重复评估之前的动作。

七、临床研究

RBMT得分低见于多种影响记忆的疾病，包括痴呆、间脑损伤、神经毒性物质中毒、乙醇（酒精）相关障碍、脑卒中、精神分裂症、颞叶切除、脑外伤。也有证据表明RBMT得分与昏迷时间和外伤后遗忘持续时间中等相关。安非他命使用者在停用药物2周后测试RBMT，可发现前瞻性记忆损害，连续使用安非他命（超过1年）与故事即时和延迟回忆进行性下降有关。轻、中度智商下降者对语言相关信息的再认（故事回忆、记姓名）有困难。

正如所料，阿尔茨海默病（AD）患者的RBMT表现差。日本Matsuda等（2002）提出识别AD的划界分是13/14（总分24分）。一个修改版RBMT（例如，合并正确命名图片的附加分，增加故事即时和延迟回忆的线索回忆）被证明对于追踪AD患者的变化具有重要作用（超过3年干预）。另外，RBMT有助于区分血管性和非血管性痴呆。Glass（1998）发现血管性痴呆患者在记约定、故事即时回忆、路线即时和延迟回忆时的表现优于非血管性痴呆。RBMT-C可用于轻度和无严重残疾的唐氏综合征的儿童和成人。

RBMT对区分各种疾病具有良好的敏感性。Perez和Godoy（1998）报道了RBMT可以识别80%的AD患者、75%主诉记忆障碍的老年人、75%的复杂部分性癫痫患者和60%的正常人，平均达72.5%的鉴别能力（WMS-R只有66.3%）。

语言表达障碍会对记姓名、故事延迟回忆、定向等项目有影响，知觉障碍则会影响路线即时和延迟回忆、面孔再认、定向和日期的表现。Wilson等（1989b）认为在评估有语言或知觉障碍患者时，这部分测验应去除，可根据其

他测验评估患者认知功能。Wilson 等报道了无语言障碍的左半球脑卒中患者和右半球脑卒中患者具有相同的测验表现，提示疾病损伤半球可能与相对保留的日常记忆能力的关系不大。

根据 RBMT 的作者所说，RBMT 不受患者焦虑和抑郁影响。然而，有证据表明，焦虑能激活监控或检查程序，可能通过干扰工作记忆能力影响测试表现，如记约定。抑郁患者的测试表现也受损。Tarbuck 报道了焦虑或抑郁恢复期的测试表现有所改善，但实验缺乏对照组，是否有混杂因素影响仍需考虑。同时，有证据表明创伤后应激障碍的儿童和青少年（11～17 岁）相比对照组会出现记忆损害表现。

有证据表明脑损伤患者在常规 RBMT 中的得分正常或接近正常，但可在 RBMT - E 中出现损害，特别是路线回忆和信件回忆。也有证据表明测试对识别轻度记忆损害敏感。Stephens 等报道了使用激素替代治疗女性的RBMT - E的得分优于未使用者，特别是故事即时和延迟回忆及信件延迟回忆项目。

八、生态/预测效度

RBMT 与临床医师所观察到的记忆问题和患者或家属主观认为的记忆障碍均有良好相关性。在考虑临床前表现时，RBMT 的得分可能与记忆障碍主诉有联系。一些学者认为相比于客观认知功能检测，RBMT 与主观的日常记忆障碍的相关度更高。但 Koltai 等报道了 RBMT 和其他认知测验如WMS - R相比，在评估暴露于神经毒性物质时患者的日常记忆功能无明显差异。

RBMT 得分与就业状态、学习新技术的能力、社区融入/社会交往能力、日常社会心理功能、日常身体状况、日常活动和生活独立能力中等相关。例如，Wilson 和他的同事报道了 RBMT 标准分低于 12 分的患者不能独立生活、有偿就业及全日制教学，而大多数 RBMT 标准分高于 12 分的患者可以从事上述活动的一项或多项。

相比于标准认知测验，RBMT 更有利于区分独立和非独立的个体，虽然这只适用于特定的人群（如痴呆患者、精神分裂症患者）。有证据表明，使用附加测验的执行功能项，特别是连线测验 B，可以提升日常适应能力的预测效果。

RBMT 也可以预测脑外伤患者记忆功能的长期预后。Wiseman 等报道了出院前的 RBMT 基线分数可预测脑外伤患者 10 年后记忆损害。同时 RBMT 基线分数也可预测精神分裂症患者社会功能的长期预后。

九、评价

RBMT是一个有效的记忆评估测验，特别是评估日常生活所需的记忆功能。RBMT适用于各个年龄段和不同文化背景的患者。RBMT具有合理的聚合效度和扩散效度，适用于多种患者，生态效度高，可能在预测日常记忆功能上相比传统的标准化记忆测验效果更好。

RBMT－II针对故事、路线等任务有即刻与延迟回忆项目，可以预计，RBMT－II能够有效识别MCI与预测MCI转化为痴呆，事实上，已经有研究验证了这个预测价值，如西班牙Bolló-Gasol的研究。

前瞻记忆测验是记忆测验中独特的部分，包括基于时间的前瞻性记忆测验与基于事件的前瞻性记忆测验。未来信件测验对早期痴呆的检测敏感性更高。但需要注意是，信件测验作为一个前瞻性记忆测验还未被普遍接受。

RBMT－II基本上没有需要握笔书写的项目，所以，RBMT－II适用于文盲与低教育人群。香港Man等(2009)开发了RBMT－II电子版，可以在线完成测验。

值得注意的是，没有指南规定故事回忆项目中叙述故事的速度。有证据证明，叙述故事的速度可以影响之后回忆的得分，叙述较慢时的回忆得分较叙述快时得分高。不同版本的测验重复评估时，不同的叙述速度可能造成对评估表现的错误推断。所以，应标准化故事叙述速度以减少错误的结论。

RBMT/RBMT－Ⅱ的部分测验有天花板效应，不适用于检测轻度记忆障碍患者，更适合检测中、重度患者。RBMT－E适用于轻度患者。RBMT虽然可以检测特殊的记忆障碍，但不是非常敏感，需要辅助更多传统实验室检查以发现记忆损害的本质。

RBMT/RBMT－Ⅱ与MMSE高度相关，意味着给予可疑痴呆的患者日常的长时间的RBMT测验意义不大，但在患者MMSE整体分数下降时RBMT各评估项目的表现可以为患者某些能力是否保留提供依据。

RBMT/RBMT－Ⅱ可应用于儿童，但在解释分数时需注意智商低于正常水平的儿童。Fraser等提供了60岁及以上老年人BRMT标准值，但高估了60～69岁老年人的表现，低估了70岁以上老年人的表现。建立RBMT－E标准值的研究样本量太少，目前使用的标准值需要调整，因为根据这个标准，25％的正常人将出现假阳性。

对于存在语言障碍或知觉障碍的70岁以下患者，推荐使用修改版

RBMT。但不推荐给存在语言障碍或知觉障碍的70岁以上患者，针对这些患者，不需要计算测验总分，只需要比较单个测验项目的得分。

修改版RBMT也适用于存在行动障碍的患者。但通过模型进行路线回忆与患者亲自在房间走路线在反应记忆能力上具有轻微不同，因为患者亲自在房间走路线时所产生的自身位置与环境的变化在模型路线回忆中是体现不出来的。

RBMT仍存在一些不足之处。首先需要更多的证据研究测验的社会心理学意义。其次，虽然测验的评估者间信度很高，但反映标准版和儿童版RBMT的内部一致性和重测信度资料不足，而扩展版RBMT则缺少这些资料。最后，反映儿童版RBMT的效度资料仍缺乏。

（陈科良　郭起浩）

第八节　Benton视觉保持测验(BVRT)

Benton视觉保持测验(Benton visual retention test，BVRT)，又名Benton测验，是Benton等于1946年编制的，到1992年，Sivan已经修订出版了第五个修订版(BVRT-5)，该测验的目的是评估受试者的视觉记忆、视知觉和视觉构造能力。

BVRT有两种测试模式，需要受试者通过画图或选择做出回答。画图模式的BVRT有3个相同难度的平行版本(C、D和E)。每个平行版本有10张图案，前两张只有1个主要的几何图形，后八张包括2个主要图形和1个更小的次要图形。

画图模式有4种常规测试方法。方法A为给受试者看图案10 s，之后立即要求受试者在白纸上默画出来。方法B与方法A相似，只是受试者只看图案5 s。方法C为要求受试者照着图案临摹下来。方法D为给受试者看图案10 s，15 s后要求受试者在白纸上默画出来。

选择模式的BVRT有两个相同难度的平行版本(F和G)，用于评估受试者的再认能力(方法M)。因为其受语言影响小，方法M也适用于不能说英语的受试者。选择模式的BVRT可用于伴或不伴运动障碍的受试者，区分受试者究竟是记忆、知觉还是画图能力障碍。

1994年Benton发表了Benton视觉形式辨别测验(Benton visual form

discrimination test，VFDT)，将 BVRT - MC 重新改编，测试者呈现每个目标图，同时呈现待选图案，让受试者选择 1 个正确答案。每个项目有 4 个多选图：1 个正确图案、1 个旋转图案、1 个碎片图案、1 个边缘不正确图案。

一、测试方法

1. *画图测试模式* 简单地说，给受试者 10 张 21.5 cm×14 cm 白纸，要求受试者在白纸上回忆(方法 A、B 和 D)或模仿(方法 C)。测试者在受试者画完后在右侧记录数字，以便确认所画图案和图案方向。

2. *选择测试模式* 方法 M 为呈现 15 张图案，包括 1～3 个几何图形，呈现 10 s。10 s 后取走图案，要求受试者在 4 个相似的选项中正确选出之前所呈现的图案。

3. *其他测试方法* 方法 O 为取走图案 15 s 后要求受试者选择。方法 P 最初用于儿童，刺激卡片和方法 M 的图案同时呈现。在方法 PR 中，受试者被需要照着图案画 10 张图(如方法 C)，然后给予方法 M 图案，并从中选择出之前所画过的图案。

每个方法的测试时间 5～10 min。

二、评分方法

通过操作手册完成评分。简要地说，画图测试模式有两项评分系统(正确分和错误分)，正确分共 10 分，每项图案只有正确(1 分)和错误(0 分)两种。具体评分原则和例子见操作手册。错误分包括对受试者表现的定量和定性分析。有 6 种错误类型：① 遗漏错误；② 变形错误；③ 重复错误；④ 旋转错误；⑤ 误给错误；⑥ 大小错误。每种错误类型还有特殊的错误亚型。另外，还要区分错误是在左侧还是右侧。错误分记录和计算见操作手册的记录表。表格中测试者可以评定每项图案是否正确，统计各种错误类型。

选择测试模式记录正确选择数(最大是 15)。

三、人口统计学影响

1. *方法 A*

(1) 年龄：年龄是预测测试表现最强烈的因子，可以解释 9%～18%的得分变化。测试得分从 6 岁开始逐渐升高，直到 14 或 15 岁达到平稳。平稳期持续至 30 余岁，从 40～50 岁开始得分又逐渐下降。显著下降从 75 岁开始。

BVRT 的变化性随年龄增加，但这种增加并不明显。

（2）性别：大多数研究发现性别对测试表现几乎无影响。

（3）教育程度/IQ：BVRT 表现与智力和教育程度具有中到高度相关（$r=0.3\sim0.7$）。教育程度的影响与年龄的影响相互联系。如教育程度高的老年人随着年龄增长得分下降的幅度较教育程度低的老年人小。在中、重度认知损害的老年人中，神经病学损害掩盖了教育程度的保护作用，以至于其 BVRT 表现下降不受年龄及教育程度的影响。

（4）职业：职业是另一个重要影响因子。Dartigues 等（1992）根据对 2720 位社区健康老年人的研究发现无论教育程度如何，BVRT 的结果与职业的相关度都高，特别是农民、家政服务业者和蓝领的记忆能力比专业人员和管理人员差 2～3 倍。

2. 方法 M

（1）年龄：一般来说，成年人的错误数少于 2 个。随着年龄的增加，测试表现逐渐下降。

（2）性别：性别对测试的影响最小。

（3）教育程度/IQ：测试表现受教育程度影响。Le Carret 等认为高教育程度受试者的表现较好，部分原因是他们在寻找目标时能使用更多的搜索策略。在儿童中，测试表现与 IQ 具有中等相关。

（4）种族：有研究发现高加索老年人和非裔美国老年人、说西班牙语和说英语的老年人之间的测试表现有差异，其中高加索老年人和说英语的老年人的得分更高。然而，在考虑高教育程度的影响后，高加索人和非裔美国人之间的差异消失了。

四、信度

1. 内部一致性　使用方法 A 时版本 C 正确分的克隆巴赫系数为 0.76，版本 D 为 0.79，版本 E 为 0.79。相对的错误分克隆巴赫系数分别为 0.71、0.82、0.80。Steck 等（1990）发现版本 C、D、E 全使用时，内部一致性系数提升至 0.91。Steck（2005）从 3 个版本的图案中选出 20 幅组成平行版本，也具有高信度。选择模式（方法 M）的折半信度为 0.76，但在儿童中（平均 10 岁），它的内部一致性较低，只有其中的 5 项能区分测试表现的好坏。

2. 重测信度和练习效应　Youngjohn 等（1992）对 17～82 岁志愿者进行了基线和 21 天的测试，发现正确分的重测信度为 0.57，错误分的重测信度为

0.53。测试是否存在练习效应仍有争议。Botwinick 等(1986)发现对 64～81 岁的老年人以 18 个月为间隔测试 4 次,得分没有明显变化。相似的是,Lezak (2004)对健康对照组于基线、6 个月和 12 个月进行测试,未发现正确分和错误分有显著差异,一致性系数分别为 0.74 和 0.77。但 Larrabee 等(1986)报道了 60～90 岁的老年人在间隔 10～13 个月后重测的得分提高了 1 分或以上。Youngjohn 等(1992)在对 17～82 岁对照组间隔 3 周的研究中也报道了类似的结果。

3. *复本信度* Sivan(1992)报道了 3 个版本(C、D 和 E)的相关系数范围为 0.79～0.84。对选择测试模式,版本 F 和 G 的复本信度较高(0.80)。

4. *评定员之间信度* 对于画图测试模式,正确分和错误分的评定员之间信度很高(>0.95)。各错误类型中,遗漏错误、重复错误和旋转错误的评估者间的一致性高(分别为 0.96、0.88 和 0.88),而误给错误和大小错误的评定员之间的一致性比较低。严格通过操作手册进行评分可以增加评定员之间信度。

五、效度

各方法间相关性:正确分和错误分的相关性高。各方法中,模仿任务(方法 C)和回忆任务(方法 D)具有中等相关。即刻回忆(方法 A)和延迟回忆(方法 D)的相关系数范围为 0.40～0.83。选择模式和画图模式的相关系数为 0.55。

因子分析显示 BVRT(方法 A)首先反映视知觉-运动因子,其次反映记忆-注意因子。一项二阶因子分析表明,综合分析其他记忆测验和日常记忆测验,BVRT 反映"警戒"和"心理活动速度"。在学习障碍的儿童和青少年中,BVRT(方法 A)与语言记忆测试(选择性回忆测试)的相关系数低(<0.25),主要反映视觉-运动能力。Lockwood(2011)针对 610 例退伍军人医院的住院患者进行研究,全部完成 BVRT、BVRT - MC、VFDT 和 WAIS - R 测试,统计发现前三者能够解释 WAIS - R 的 4 个因子的 81%。

然而,Moses(1986)报道认为方法 A 和选择模式都首先反映记忆能力,其次反映注意广度和知觉分析能力;模仿任务(版本 C)则首先反映注意广度和知觉分析能力。一项 162 名神经精神患者的验证研究明确了 BVRT 模仿任务和回忆任务具有不同的因子组成。Dougherty 等(2003)发现在分裂性行为障碍的青少年中,其冲动行为(如在连续测试中出现替代错误)影响 BVRT 测试

表现，说明执行控制能力也参与BVRT表现。

选择模式中一些项目可以不需要看到原图，仅通过解决独特问题的方法得到正确答案。因此，测试的正确性更依赖于受试者回答时的策略性而不是视觉记忆。Franzen(1989)建议测试者可在测试结束后与受试者面谈来确定受试者的策略。

六、临床评价

多项研究评估了不同疾病患者的BVRT表现。这些研究表明标准版测试(方法A)对识别神经行为障碍敏感度好，但诊断性不高。Steck等报道，即使在30个项目的测试中，未发现抑郁、精神分裂症、乙醇(酒精)中毒和脑损伤患者的得分和错误类型有显著差异，虽然他们的错误分都比正常对照组高。

BVRT(方法A)损害见于多种存在影响记忆状态的情况，包括脑外伤患者、复发缓解型多发性硬化患者、亨廷顿病基因携带者、多种药物滥用者、未激素替代治疗的绝经后妇女、携带*ApoE*基因和MRI显示脑萎缩的老年人。

皮质下痴呆和皮质性痴呆患者BVRT(方法A)的表现差。即使是轻度痴呆，BVRT也具有高敏感度。Robinson-Whelen(1992)报道了轻度和中度痴呆患者与正常人BVRT(方法A和方法C)的表现具有显著差异。在两种方法中，患者和对照组的各种错误类型均有显著差异，其中遗漏错误显著升高。Storandt等(1986)也报道了轻度AD患者较年龄匹配的正常人的方法C测试错误数显著增高(平均3.3，标准差5.1)；2.5年后随访，患者组的错误数较基线显著增高(平均13.5，标准差11.7)，而正常对照组则无明显变化。

Zanini(2014)报道BVRT能够有效区别巴西的正常老人与AD患者。有证据表明BVRT(方法A)可以提前10年以上预测AD的发生。Kawas等(2003)报道了BVRT错误分6分或以上的受试者15年以后发展为AD的风险是错误分0～5分受试者的2倍。单种错误类型与AD风险无关。Swan等(1996)发现在回归健康状况后，如肿瘤、心血管病、收缩期高血压和血脂水平，BVRT是老年人死亡率高低的强烈预测因素。

临床推测右半球后部区域损害的患者在画图模式BVRT中损害最重，但对该推测的可靠性仍存在争议。例如，DeRenzi等(1977)和Vakil等(1989)都发现右半球和左半球损害患者方法A测试中的表现都较正常组差，但相互间无显著差别。右半球损害患者的方法D测试的正确分较左半球损害患者低，但两者的错误分无差异。Vakil等(1989)则认为两个分数都存在差异。Mann

等(1989)发现多发性硬化患者 MRI 脑损害体积与 BVRT 的损害程度相关。双侧半球损害的患者可见 BVRT 损害,但左侧相同区域损害患者的 BVRT 表现更差。另外,方法 C 和方法 M 均不能区分受试者左侧或右侧半球损害。

Arnklsson(1993)针对儿童的研究显示标准方法(方法 A)能有效区分留级和正常 5 年级学生。Snow(1998)也发现该测试能有效区分不同学习障碍的儿童。

Baum 等(1996)对多种神经心理学测试进行了相关分析。BVRT 具有 0.85记忆权重和 0.69 模仿权重,提示在 AD 患者中具有良好的生态效度。BVRT 表现差也与 AD 患者的决策能力差和儿童时期严重脑外伤患者的职业结果损伤有关。然而,在正常学生中,BVRT 的模仿和记忆分与年级无关。

BVRT 也用于评估脑外伤和乙醇(酒精)中毒患者的记忆训练和精神分裂症认知和沟通训练。

模拟人和诉讼当事人在诈病时会产生比脑外伤患者更少的正确回答和更多的错误。特别是,他们会产生比脑损伤患者、抑郁患者或躯体障碍患者更多的变形错误。

BVRT 具有多项优点:测试时间短,评分标准精确,评定员之间信度高,内部一致性合格,各版本实用性高。另外,测试者可以通过选择、模仿、回忆图形等不同的方法,检测受试者的知觉、运动和记忆能力。

各种临床研究已经证明 BVRT 在评估年龄相关认知下降、痴呆、头颅外伤和学习障碍中的敏感度较高,虽然该测试对右半球损害的敏感度不是那么令人信服。另外,错误类型能有效地发现忽视。

使用者必须注意标准值的限制,该测试不仅用于说英语的人群,还用于其他国家如中国、埃及、印度、委内瑞拉。标准值的应用范围因测试方法不同而不同,使用者需要根据不同的方法标准值使用年龄和教育范围选择合适的测试方法。尤其要注意的是,简单的教育成就(最高学历)不能完全反映不同种族教育经历的不一致。其中一种避免错误的方法是根据阅读再认进行调整。

方法 A 和方法 M 在高教育程度的青年和中年人里容易出现天花板效应,在解释结果时需要注意。对重测信度存在争议,并只限于成年人。另外,文化影响只在选择测试模式中被系统地检测了。最后,虽然测试想要反映受试者的非语言记忆,但是一些几何图形是可以言语化的,一些选择模式的项目可以通过策略回答。而且,相比视觉记忆,模仿测试模式更与视知觉-运动能力相关。

(陈科良　郭起浩)

第五章

语言的评估

语言交流在人类社交中起核心作用。常用的语言评估量表包括各种版本的失语筛查测验、Boston 命名测验(BNT)、词语流畅性测验(VFT)、双听测验(dichotic listening test)、表达词汇测验(expressive vocabulary test, EVT)、Peabody 图-词测验(Peabody picture vocabulary test, PPVT)、成人阅读测验(national adult reading test, NART)、标记测验(token test, TT)和韦氏智力测验的语言分测验。本节没有介绍各种版本的失语检查方法,而是结合痴呆患者临床评估的需要,介绍爱丁堡功能性交流能力检查法(Edinburgh functional communication profile,EFCP),它尽管是为失语患者编制的,还没有在痴呆患者使用的报道,但是,笔者认为它采用的观察模式有助于临床上与痴呆患者的沟通与评估,有兴趣的医生不妨试用。

第一节 词语流畅性测验(VFT)

词语流畅性测验(verbal fluency test, VFT)又被称为受控词语联想测验(controlled oral word association, COWA)。词语流畅性的说法因容易被误认为是人际交流中的口语能力或句子连续性而在某些论文中采用 COWA 的名称。VFT 要求受试者就某一范畴在有限的时间(通常为 1 min)内列举尽可能多的例子,例如,请你说出所有你记得的花的名字,你可以说玫瑰、菊花、剑兰等。常用的范畴有动物、水果、蔬菜、服装、交通工具、姓氏、城市名、超市商品、家庭用品和 F 或 A 开头的单词。最后一种又称为字母流畅性测验或音位流畅性测验,在英语国家中很常用。如果将语义联想和语音联想结合起来,可

以要求列举“C开头的动物”。词语书写流畅性测验是要求3 min内写下尽可能多的F或A或S开头的单词。假如受试者停顿15 s，应该重复一下指导语。要记录所有的回答。

一、类型

1. 范畴流畅性　这是最常用的一种类型，列举动物的例子（animal fluency test，简称AFT）、列举水果的例子、列举蔬菜的例子、列举家庭用品的例子、列举超市商品的名称、列举服装的名称、列举交通工具名称、列举中国人的姓氏、列举你所知道的城市名称，等等。因为文盲老人往往不明白“动物”是什么意思，所以，尽管“列举动物的例子”是ADNI或CERAD等成套测验的组成部分，它在低教育人群的适用情况有一定的局限性。

2. 动作流畅性　如能在厨房里做什么事情？能在教室里做什么事情？能在公园里做什么事情？能在田里做什么事情？作为比较，可以同时问：你在厨房里能找到哪些东西？学校教室里有哪些东西？公园里有哪些东西？田里有哪些东西？

3. 观念流畅性　如分别列举“白纸”和“瓶子”有哪些用途。回答有两类：传统回答如“白纸”用来写字等，发散回答如“白纸”用来折飞机等。

4. 图案流畅性　常用的有两种：一种是5 min内产生尽可能多的不同的抽象图形（design fluency test），另一种是5点测验（five-point test，FPT），要求受试者用4条线将5个点连接起来。限时3 min。

5. 音位流畅性　又称为FAS测验，列举F、A或S开头的单词，是反映额叶执行功能的测验，可能与这个测验功能相近的是：列举包含汉字“高”、“水”或“发”的词语、俗语和成语；交替列举流畅性测验：如动物名称与城市名称交替列举（牛-广州-猫-杭州-兔子-北京）。

在中文语境下，目前还没有难度和脑定位方面与FAS完全对应的测验。笔者曾试用以下方式，但是没有成功：① 要求列举“大”或“发”开头的词语，文化程度的相关性高，不严格的回答多（比如“大”开头的词语有大学、大河、大嫂等，但不少受试者喜欢回答大手、大鼻、大桌子等眼前事物加一个“大”字，对错难辨）。② 要求受试者列举尽可能多的木字旁、提手旁或三点水旁的汉字，这种方法不仅教育程度影响非常大，而且列举总数偏少，我们调查105例初中教育程度老人、114例高中教育程度老人、120例大学教育程度的正常老人，偏旁流畅性测验总分（2 min分别列举口字旁和草字头旁汉字之和）分别为8.4、9.4

和 11.4 个，远低于原版中 A 开头的单词数量。③ 直接采用拼音为 F、A 或 S 开头的汉字，由于老人不懂拼音的比例非常高，尤其是南方老人，平时以讲方言为主，不容易实施。

二、评分

1. 评分指标

（1）正确数：想象的或神话中出现的动物如龙、麒麟应该算正确。

（2）错误数：包括专有名词（如赤兔、千里马）、错误的、不属于该范畴的和重复出现的例子。英语背景的错误有 4 种形式：重复（持续言语）、插入（不属于该范畴的例子）、错语症和拼字错误。中文背景的错误形式有显然差别，没有“拼字错误”。别称应该作为重复，计算其中一个，如同时列举老虎和大虫，应计算为 1 分。

（3）不同时间段列举的正确数和错误数。如前 15 s 和最后 15 s 列举的例子占的比例是多少。

（4）归类程度：这是反映语义策略运用的能力，如列举“家中用品”时，厨具类、家电类和家具类分别是多少，每个类别的串联程度（semantic cluster），如列举“动物”时，哺乳类、鱼类、鸟类、昆虫类分别是多少，按照类别列举的例子占的比例是多少。不同类别的转移次数也是进一步细致分析时的常用指标。

（5）语义串联数和类别之间转换次数：如列举超市商品时，有“苹果、香蕉、可乐、纯水、梨子、猪肉、草纸”，正确数为 7、错误数为 0、重复数为 0、串联数为 2、类别之间转换次数为 4。

2. 划界分的演变　1991 年上海社区老年性痴呆流行病学调查时采用了言语流畅性测验，其中 AFT 的划界分文盲 6 分、小学教育程度 7 分、初中教育程度及以上 8 分，敏感性 77.8%，特异性 76.7%。2004～2006 年复旦大学附属华山医院神经心理室的调查资料：来源于社区的正常老人 511 名，AFT 总正确数 16.3±4.4 个，来源于本院神经内科记忆障碍门诊的轻度 AD 患者 67 名，AFT 总正确数 10.4±3.5 个，划界分初中组 11 分，高中组 12 分，大学组 13 分，敏感性 81%，特异性 81%。2009～2011 年 AFT 的正确数 18.2±4.0 个。即，在年龄、教育基本相同的情况下，正常老人 AFT 总分 1991 年约 13 分、2005 年约 16 分、2010 年约 18 分。呈现弗林效应（Flynn effect），弗林效应指智商测试的结果逐年增加的现象。这个纵向资料也表明标准化的神经心理

测验数据应该及时更新，否则容易出现假阴性。

三、评价

言语流畅性测验简单实用，是各种成套测验的一部分，如认知能力筛查量表(CASI)、长谷川痴呆量表(HDS)、七分钟痴呆筛查测验、Dem tect(Kalbe，2004)、Mattis 痴呆评定量表、剑桥老年精神状态检查法和美国 CERAD 成套神经心理测验均包括言语流畅性测验。

不管是正常人还是轻度痴呆患者，在限定的 1 min 内以前 15 s 列举的例子最多。前 15 s 产生的词汇更多的是反映大脑的一个自动加工过程，是速度和活动的指数；后 45 s 产生的词汇反映了工作记忆、定势转换和执行控制，可以通过对受试者产生词汇的策略进行仔细分析而获得。诸如动物、水果的范畴提示可能依赖左颞顶后部的功能，而语音提示的词汇表的生成(FAS 测验)则与左额背外侧完整性有关。

范畴流畅性测验要比字母流畅性测验容易，尤其是前 15 s 最明显，列举动物的例子是 C(或 F、L)字母开头的单词数的 2 倍，这是因为这两种流畅性测验涉及的层次组织不同，列举动物的例子只要从哺乳类、鱼类、鸟类等几大亚群中提取具体的名称，列举字母开头的单词需要更多的语义范畴(甚至每个单词属于不同的语义范畴)，所以，针对范畴流畅性测验制定的划界分要比字母流畅性测验的高。由于汉字的偏旁有意义提示作用，列举木字旁的字可能与列举植物名称的复杂性相似，范畴流畅性和字母流畅性是否存在差异还未见报道。

左侧与双侧额叶损伤，如前交通动脉瘤破裂延及额叶、多发性硬化的胼胝体前部萎缩、左丘脑后结节切除术后、左颞叶切除术后，COWA 的表现明显下降。针对正常志愿者的 PET 研究发现，受试者完成 COWA 时，右背外侧前额皮质和内侧前额的血流量增加。AD 患者字母流畅性的损害比范畴流畅性重。如果给予亚群线索，如家养的动物、野生的动物，帕金森病和亨廷顿病所致痴呆患者可以明显提高列举的数量，而 AD 患者没有改善。额颞叶痴呆(FTD)在 COWA 的表现显著低于严重度匹配的 AD 组的表现。由此可见，COWA 反映的是额叶执行功能。

有时门诊为了节约时间，只给患者 15 s 或 30 s 进行列举，则不能达到准确识别轻度认知损害(MCI)的目标。笔者统计前 15 s 的列举数，在年龄、教育匹配的正常老人和 MCI 两组之间没有显著差别，前 30 s 和前 45 s 的鉴别力也不

如做完 1 min 的结果，故尽管流畅性测验耗时短，但针对轻度患者，最好不要少于 1 min(表 5 - 1 - 1)。

表 5 - 1 - 1 动物流畅性分时指标比较

指 标	正常对照组(n=171)	MCI 组(n=83)	t 值	P 值
1～15 s	7.6(1.9)	7.2(1.9)	1.566	0.119
16～30 s	4.6(1.9)	3.3(1.8)	4.922	0.000
31～45 s	3.4(1.7)	2.0(1.4)	6.257	0.000
46～60 s	2.6(1.5)	1.7(1.4)	4.414	0.000
总正确数	18.2(4.0)	14.3(3.9)	7.384	0.000

资料来源：笔者尚未发表的数据。

在国内，大部分针对社区老人的痴呆流行病学调查包含言语流畅性测验。如 Mok 针对没有受过正式教育的老人样本，FVT(列举动物、水果和蔬菜的总数)筛查痴呆的敏感性是 86.8%，特异性是 93.4%。VFT 受到受试者的年龄、性别、教育程度、智力、记忆力的影响。相对而言，语义流畅性测验依赖颞叶结构，反映记忆功能，而音位流畅性测验依赖额叶结构，反映执行功能。左侧与双侧额叶损伤，如前交通动脉瘤破裂延及额叶、多发性硬化的胼胝体前部萎缩、左丘脑后结节切除术后、左颞叶切除术后，VFT 的表现明显下降。针对正常志愿者的 PET 研究发现，受试者完成音位流畅性测验时右背外侧前额皮质和内侧前额的血流量增加。如果给予亚群线索，如家养的动物、野生的动物，帕金森病和亨廷顿病所致痴呆患者可以明显提高列举动物的数量，而 AD 患者没有改善。额颞叶痴呆在 VFT 的表现显著低于重度匹配的 AD 组的表现(表 5 - 1 - 2)。

表 5 - 1 - 2 AD 与 FTD 两种痴呆类型的流畅性测验的差异

项 目	AD	FTD
病理易感部位	海马/颞叶	额叶皮质
敏感的流畅性类型	语义流畅性测验	音位流畅性测验
词性	列举名词	列举动词
内部信息加工	语义串联数(cluster)	类别之间转换次数(switch)

尽管汉字的偏旁有意义提示作用，列举木字旁的字仍然少于列举植物名称的例子。有趣的是，在列举动物名称时，去除生肖动物名称，以“非生肖动物名称”作为判断指标，识别 AD 的敏感性比总正确数指标提高了 4 个百分比(特异性相同)。

第二节　波士顿命名测验(BNT)

失语症患者和其他神经疾病患者的命名障碍非常普遍，目前最常用的是Kaplan编制的波士顿命名测验(Boston naming test，BNT)，BNT是用常见物品的黑白图画评估视觉命名功能(图5-2-1)。1978年Kaplan等人最初编制了这个常用测验，该版本有85项，1983年Kaplan等人将其修订为60项。目前通用版本(BNT-2)是保留了60项的版本，还包括一个包含15项的简短版本及一个多选版本。被称作BNT-2的版本是增加了难度的物体线条图画，范围从简单的高频名词到少见名称，要求对从易到难排列的60幅线条图进行自发命名和线索命名。

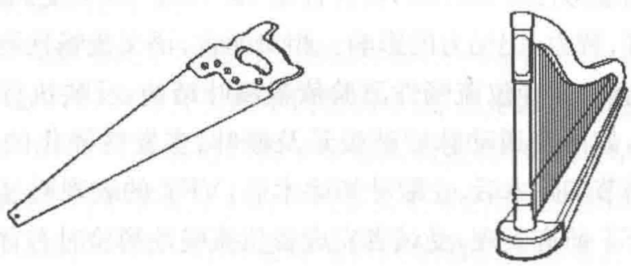


图5-2-1　BNT图片举例

一、指导语

在使用长版(BNT-2、BNT-60)时，对于儿童和失语症患者，以条目1开始，到连续8个错误时终止(这个测试之前的版本为连续6个错误时终止)，对于其他所有成人，从条目30(口琴)开始，如果接下来的任何8个条目错误，则转到条目29开始反向继续进行，直到8个连续条目能够无须提示通过(如无备选答案或语音提示)。之后，返回到往前的、难度逐步增加的方向继续测试，当受试者出现连续8个错误时终止测试。

如果在20 s内正确命名该图形则给分。只有当受试者明显误解某幅图片时，才被告知图片像其他东西和提供括号里的语义线索。无论是自发的还是语义提示，每次提示后仍回答错误则给予语音提示。例如：受试者将“蘑菇”回应为“雨伞”，就提示他/她“是吃的东西”，并且再给20 s命名该图片。如果受试者在20 s内正确命名该条目，核对线索提示后的正确数，如果

给予线索提示,患者无法在 20 s 内正确命名该图片,就给予语音提示,如强调该条目首字母为"m"。又比如,有一幅图是"飞镖",受试者回答正确,接着做下一题,如果回答不正确,比如是"羽毛",可以问"它还有别的名称吗?",如果回答错误,比如是"毛笔"或不能回答,就给予语义线索:"这是用来投掷的东西",假如在 20 s 内依然不能回答,就给予语音线索:"这个名称是以 d 开头的单词"。

终止和开始测试的规则的制订是以节省大量时间并对受试者无明显影响为前提条件。而终止测试的规则没有在测试指南中明确指明。语音提示后的正确回应不计入总分。而指南并未明确指出语音提示是否应该用于决定测试终止。Weintraub(2003)做了严格解释,包括在计错误数时对语音提示的正确回应数。Ferman 等人 1998 年报道了使用"宽松的"终止方法(在计 6 个连续错误数时不包括语音提示正确回应数)导致了 655 名正常老年人评分变化了 3%,140 名 AD 患者评分变化了 31%。在正常受试者中,差异评分最常见于≥80 岁人群,且得分差异上升至 16 分。因为终止规则的不同解释可能会改变测试得分,因此在公布常模数据时清楚的描述规则是很重要的,比如 Ferman 等人 1998 年提供的常模数据是用 6 个而非 8 个条目为终止规则收集的。

二、中文版 BNT

1. 中文版 BNT 的修订　作者用的中文版 BNT 共 30 幅图片(来源于全本 BNT),第一步自发命名和第二步语义线索命名与原版相同,第三步改为选择题,由正确答案、形态相似名称和同类物品名称组成的 3 个名词随机呈现,请受试者选择一个,如"这是标靶、飞镖、火箭三者中的哪一个?" 耗时10～20 min。

2. 不同文化背景 BNT 比较　作者在使用 BNT 中文版时采用从第一项到最后一项全部都做的方法,这是因为中文版按照原版次序排列,并且不同地区的难易次序不同。从表 5－2－1 可以看出正常老人的 BNT 表现与文化背景的关系。与中国香港受试者比较,上海受试者对于海马和飞镖的识别比较差;而对冬菇和手风琴的识别比较好。与美国受试者比较,上海受试者对于海马、飞镖、竖琴和冰屋的识别比较差;而对算盘、漏斗和手风琴的识别比较好。总体来说,中国香港受试者的表现介于上海受试者和美国受试者之间,即上海受试者与美国受试者的差异多于与中国香港受试者的差异。

表 5-2-1 中国上海、香港和美国正常老人受试者的表现比较

图片名称	中国上海受试者回答正确率(%)	中国上海受试者的排名	中国香港受试者的排名	美国受试者的排名
树	100	1	1	1
笔	100	2	2	2
剪刀	100	3	3	3
花	100	4	4	4
衣架	100	5	5	8
骆驼	100	6	7	10
羽毛球拍	100	7	10	11
算盘	100	8	8	30
锯子	99	9	6	5
冬菇	99	10	25	7
漏斗	97	11	9	23
手风琴	97	12	24	24
扫把	96	13	11	6
轮椅	95	14	12	9
听诊器	94	15	18	21
圆规	93	16	13	25
扶手电梯	91	17	14	19
蜗牛	86	18	15	12
钳	84	19	20	27
口琴	81	20	19	15
三脚架	81	21	26	26
犀牛	74	22	23	16
花棚	66	23	28	28
仙人掌	65	24	17	18
金字塔	62	25	21	22
量角器	58	26	29	29
竖琴	54	27	27	20
飞镖	49	28	21	14
海马	45	29	16	13
冰屋	8	30	30	17

资料来源：笔者未发表的数据。

三、不同版本的 BNT 介绍

1. 短版(Mack SF4)　短版是为了缩短测试时间编制的。Mack 等人 1992 年编制了一个 15 幅的版本,众所周知,Mack SF4 被 BNT-4 的作者采用。这个简短版本在图片和回答上都优于标准 60 项的版本。AD 联合登记组织(CERAD,Morris 等 1989 年)编制的另一个 15 条的版本目前十分常用。这些条目被严格的划分为 3 组,每组 5 条,代表了英语高频(容易命名)、中频、低频词(很难命名)的物体。Mack SF4 版本常常以条目 1 开始,8 个连续错误终止。

2. 多选版本　在测试完成后,受试者返回到经语音提示仍命名错误的条目,给出该条目的图片和 4 个备选答案,测试者阅读每个选项,让受试者标示正确选项。

3. 实用 30/15 版本　Williams 等人 1989 年编制的 30 条目版本也是目前很常用的版本。该测试已被证明有很好的可靠性和高精度,并与 BNT60 高度相关。

4. 其他版本　Graves 等在 2004 年使用项目反应理论编制了两个新的 15 项和 30 项目短版测试,它们有很高的内部可靠性(15 条目版本内部可信度为 0.84,30 条目版本内部可信度为 0.90),与长版有很高的相关性,在鉴别正常衰老、AD 及血管性痴呆患者方面与采用完整 BNT 做出的分类是一致的。作者还编制了实用版本:检查者给出 15 条目版本,如果得分在具体可信区间以外则终止测试,否则继续完成 30 条目测试。比如,检查者首先执行 15 条目版本,如果得分≥12 分,就给予患者另外 15 条目得分;如果得分≤3 分,检查者终止测试,给予目前已得分数。除这两种情况之外,检查者执行剩余条目以完成他们的 30 条目测试。

不同版本 BNT 的项目内容见表 5-2-2。

表 5-2-2　5 个版本 BNT 的项目内容

MackSF4 BNT-15	CERAD BNT-15	Williams 等 BNT-30	Graves 等 BNT-30	Graves 等 BNT-15
4 房屋	1 床	11 直升机	13 章鱼	19 长凳
7 梳子	2 树	13 章鱼	17 骆驼	24 海马
10 牙刷	4 房屋	18 面具	19 椒盐卷饼	31 犀牛
13 章鱼	5 哨子	19 椒盐卷饼	21 球拍	32 橡子

（续表）

MackSF4 BNT - 15	CERAD BNT - 15	Williams 等 BNT - 30	Graves 等 BNT - 30	Graves 等 BNT - 15
20 长凳	8 花	23 火山	22 蜗牛	35 多米诺骨牌
23 火山	10 牙刷	24 海马	23 火山	41 鹈鹕
26 独木舟	17 骆驼	28 花环	24 海马	42 听诊器
29 海狸	18 面具	30 口琴	25 飞镖	44 枪口/炮口
36 仙人掌	23 火山	31 犀牛	28 花环	45 独角兽
39 吊床	26 独木舟	32 橡子	29 海狸	48 手风琴
42 听诊器	30 口琴	33 冰屋	31 犀牛	53 书卷/画卷
45 独角兽	35 多米诺骨牌	34 桩子/高跷	32 橡子	54 钳子
52 三脚架	39 吊床	35 多米诺骨牌	33 冰屋	55 狮身人面像
55 狮身人面像	46 漏斗	37 自动扶梯	34 桩子/高跷	57 花棚
58 调色板	54 钳子	39 吊床	35 多米诺骨牌	58 调色板
		40 门环	36 仙人掌	
		41 鹈鹕	38 竖琴	
		43 金字塔	41 鹈鹕	
		44 枪口/炮口	42 听诊器	
		48 套索	43 金字塔	
		49 芦笋	44 枪口/炮口	
		50 罗盘	45 独角兽	
		51 弹簧锁	48 套索	
		52 三脚架	53 书卷/画卷	
		53 书卷/画卷	54 钳子	
		54 钳子	55 狮身人面像	
		55 狮身人面像	57 花棚	
		57 花棚	58 调色板	
		58 调色板	59 量角器	
		60 算盘	60 算盘	

备注：名称前的标号指发表的 60 条目 BNT 的条目号。

四、评分

评分包括正确数和错误数。错误的分析方法见表 5-2-3。

原版的分析指标包括：① 无线索时正确回答数；② 线索提示正确数；③ 线索提示后正确回答百分比；④ 语音提示正确数；⑤ 语音提示后正确回答百分比。

表 5-2-3 回答的错误类型

	错误类型（汉语）	错误类型（英语）	代码	定义(反应项目是目标项目的)	举例（目标词-冰箱）
语义错误					
1	同位错误	coordinate	co	同一语义范畴成员	电视
2	上位错误	superordinate	s	上位概念	电器
3	联想错误	associative	a	联想关系	牛肉
4	定义（描述性错误）	definitions (description)	d	定义	一种可以制冷的家电
非语义错误					
5	赘语	circumlocutions	ci	围绕一个说不出的词，不停地说出一串无意义的，或代以虚词	冬天湖面水出现的一层东西，……
6	视觉相似性错误	visual	v	视觉相似	书
7	语音相似性错误（普通话）	phonological	ph	语音相似（声母、韵母或声调错误）	平凉
8	方言发音且错误				
9	字形相似性错误	orthographical	o	字形相似	水箱
10	持续性错误	preservative	pr	上一个项目	
11	组词错误	word-forming	w	其中的部分词素/或音素	冰
12	词性错误	grammatical	g	不同词性的词	制冷
13	非词错误（新语错误）	nologism	n	非词或新创造的词	里中
14	完全无关	unrelated	u	无关	牙刷
15	声调错误				
混合性错误					
16	混合性错误Ⅰ	mixed	m1	同位错误＋视觉相似	
17	混合性错误Ⅱ		m2	同位错误＋语音	烤箱
18	混合性错误Ⅲ		m3	同位错误＋字形	
其他					
19	无法分类的错误				
20	无反应、不知道	don't know	d	无反应	

注：如果回答同时是“同位错误”和“视觉相似性错误”，如“老虎”命名为“猫”应该记录为 16。

中文版的分析指标为：① 无线索时正确回答数；② 线索提示正确数；③ 线索提示后正确回答百分比；④ 选择命名正确数；⑤ 选择命名正确回答百分比。

BNT 中某些被认为是不正确的回答实际上可能是某些地区常用的同义词。AD 患者会使用更多的同义词，对这些同义词的得分判断可能导致评分的细微差异，但这些差异不可能改变定量解释（quantitative interpretations），然而，临床医生应该小心避免将同义词列为错误。专家建议遇到同义词回答时顺带问一句："这个物品还有其他名称吗？"

种族、涵化（acculturation）水平、地区差异、语言背景等都对 BNT 得分有影响。所以，各地区需要建立自己语言文化背景的常模资料。不同研究采用的 BNT 版本的项目内容与数量有差异，所以，无法对各种常模数据一一介绍。在笔者采用 30 项版本对年龄、教育程度匹配的 100 名社区正常老人、38 名 MCI、34 名轻度 AD 和 38 名中度 AD 患者的调查表明，正常老人组、MCI 组、轻度 AD 组和中度 AD 组自发命名得分分别为 24.9±3.0 分、20.9±3.6 分、18.7±4.0 分和 15.7±4.2 分。以 BNT 自发命名≤22 分作为划界分，特异性均为 81%时，识别 MCI 的敏感性为 61%，识别轻度 AD 的敏感性为 79%，识别中度 AD 的敏感性为 95%。当然，有必要制定不同年龄、不同教育程度的划界分。

五、评价

正常老人的自发命名能力受到其年龄和教育程度的影响是显而易见的。性别也有影响，而且是女性的 BNT 得分低于男性，这与女性语言理解、表达和记忆的能力优于男性的一般印象不同。在 Randolph 针对 719 名美国社区正常老人的研究中，采用 BNT 60 项版评估，有 18 个项目是男性明显优于女性，仅 4 项是女性明显优于男性。所以，"女性的 BNT 得分低于男性"可能与 BNT 的项目选择有利于男性有关。

BNT 是一种常见的视觉对偶命名测验。也许除外 COWA，没有语言功能测试像 BNT 在神经心理学测试中如此常用。事实上，最近神经心理学家进行的评估实践调查显示，BNT 在使用频率上排名第四（Rabin 等，2005）。大多数健康成年人 BNT 测试都可得高分，这就是说，此测试不能很好地区别高分人群，所以，对于那些认知减退或损伤极为轻微的人群必须谨慎判断。BNT 对于检测轻度 AD、失语、皮质下疾病（如多发性硬化和帕金森病）均较敏感。AD 患者的 BNT 的低分预示病程进展将更快速。

针对 AD 患者的 MRI 研究中发现 BNT 总分与颞叶、海马和海马旁回的容积有显著相关性；PET 研究发现 BNT 总分与左颞叶代谢有关，然而，对左半球为优势半球的患者行局限的左颞叶前部切除术和右颞叶前部切除术，

BNT 的表现未受影响。

由于BNT 部分图片(如冰屋、竖琴、海马)与文化背景密切相关,对于中国老人显得难度过高,为了避免文化背景偏差,笔者曾经编制“华山命名测验(Huashan naming test, HNT)”, HNT 选择的是日常生活中的常见的动植物、工具及家庭用品的彩色实物图片(图 5-2-2),正常老人的平均回答正确率高达 86%,与 MMSE 一样有一定的天花板效应。在总分 30 分中,以自发命名正确数≤26 分作为划界分,HNT 识别轻度 AD 的敏感性和特异性约为 82%,并没有明显优于 BNT-30 项版本。考虑到研究中需要进行国际比较,笔者首先推荐的命名测验还是BNT,而在农村文盲比例比较高的地区可以采用 HNT 协助诊断与随访。

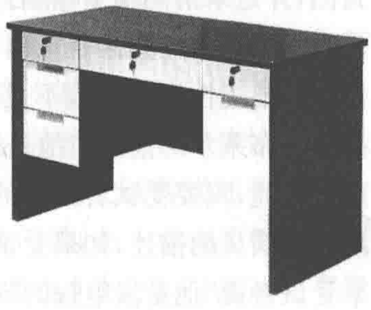


图 5-2-2 HNT 图片举例

附件 5-1 两个难度相同的 HNT 版本项目(按照由易至难顺序排列)

HNT-Ⅰ:黄瓜、剪刀、猪、电话机、香蕉、算盘、衣架、针筒、绵羊、夹子、轮船、土豆、蘑菇、马、孔雀、叉子、磅秤、石榴、电吹风、放大镜、电视机、扳手、长颈鹿、书包、直升机、订书机、火车、蜗牛、猕猴桃、轮椅共 30 项。

HNT-Ⅱ:茶壶、汽车、书桌、眼镜、茄子、萝卜、菠萝、青菜、灯泡、坦克、猫头鹰、葡萄、钳子、沙发、大象、望远镜、台灯、钢笔、血压计、花菜、衣夹、扶手电梯、钥匙、照相机、梳子、企鹅、草莓、听诊器、计算器、台历共 30 项。

资料来源:林春颖,郭起浩,赵倩华,等. 华山命名测验的编制与效度分析. 中华行为医学与脑科学杂志,2011,20(2):183-186.

第三节 名人面孔命名测验(FFT)

一、名人面孔测验-75

名人面孔测验-75(famous faces test-75, FFT-75)的测试材料包括 75 张与受试者生活在同时代著名人物的照片,他们有的还健在,如:托尼·布莱尔、大卫·贝克汉姆,而有些人已过世,如:温斯顿·丘吉尔、戴安娜。为每个著名人物选择一个同性别、同年龄、总体外貌相似的普通人面孔以减少名人面

孔"过分突出"的可能性。所有的面孔都有相同格式(11 cm×13 cm,黑白肖像照片)。每个著名人物面孔与相匹配的普通人物面孔并排在一张 A4 纸的两边,且著名人物面孔的放置位置应左右均等。

(一) 指导语

向受试者依次展示每组照片,让其指出著名人物,如某受试者无法识别可猜测,并记录猜测结果,然后让受试者提供识别信息,并尽可能地说出其姓名,只有说出全名才算命名正确,检查者不提供反馈信息。因语义性痴呆患者可使用的描述性词语严重不足,故相关信息识别可使用较宽松的标准。如:对托尼·布莱尔的描述,"他是最高领导人"和"他是前任首相"都被认为是正确的。因此,只要受试者描述的识别信息正确,即使使用的术语不对也算正确。对电影演员的描述,如果受试者说"他出演了很多老电影",就判断为正确,如果受试者说"他是演电视的吗?"则判断为错误。检查者不予回馈。

(二) 评分

该著名人物命名测验包括 75 张名人照片,评分规则如下:① 命名:无线索提示的名人命名;② 识别:提供的面孔描述、识别信息;③ 熟悉程度:基于著名人物面孔的选择判断;④ 主观熟悉程度:受试者在对著名人物面孔进行选择判断时是根据猜测还是熟悉。

(三) 评价

语义性痴呆(SD)患者比 AD 患者的命名、识别著名人物面孔正确数更少。

二、名人面孔测验-50

该版本为 Hodges 和 Ward 在 1989 年编制的改良版的著名人物面孔测试。测试选取 50 个著名公众人物的照片,要求入选人物在一个限定时期仍有知名度,这个限定时期可能仅仅是 20 世纪 40 年代到 20 世纪 80 年代的某个年代。要找到在某个年代很著名的人物的确很难,但应尽可能地尝试选取这样的名人。每个年代选取 10 张照片,这些照片包括政客和政治家、警察、戏剧和电视电影明星以及运动员,所选照片均为黑白照片,所选著名人物的详细信息应备注在附录中。

每张名人照片有 3 张普通人物照片做陪衬,并且年龄、性别、地区与名人照片相同。这些普通人物照片选取范围广泛。出于管理目的,每张名人照片都为 2×2 排列规格,为了整体设计平衡,每张名人照片的位置都是随机的,因此,该测试包括 50 组,目标照片的排列顺序为每 5 张名人照片为一组,每组中

的5张照片来自不同年代。

(一) 指导语

本测试的每张照片有识别、命名、确认3个潜在的部分组成。第一部分"识别",为受试者提供4张一组的各组照片,指导语如下:"4张图片中只有1张是著名人物,请指出你认为是著名人物的那张照片。"若回答不正确,为受试者指出正确答案。第二部分"命名",紧接着问受试者该著名人物的名字,只有回答全名才算正确。如回答正确就进行下一条目,若回答不正确,或部分正确,如:只知道姓氏,受试者将进行该条目的第三部分确认测试。若受试者只能说出该人物的姓或名,那么,检查者告知其全名,并让受试者确认该著名人物的身份。对于不能命名的人物面孔的身份确认,鼓励受试者描述著名人物的细节。测试者不可直接提供线索,但可针对细节问题,进一步使用标准探测语句以获取受试者更具体的描述。如:受试者回答为"政治家",测试者可问"他的职务是什么,代表何政党?";受试者回答为"演员"可问"他/她出演过何电影或电视剧等"。正确回答计分应包括特定识别信息,如"格伦达·杰克逊,她是女演员,目前为工党议员"等。错误回答应包括身份属性表述错误或模糊,如"著名影星、外国政治家等"。只有当检查者认为受试者:① 提供了有关著名人物面孔充足的专属识别信息;② 只能提供模糊的识别信息;③ 根本什么都不知道,这三种情况时,才认为测试完毕。

(二) 评分

评分包括总的正确数、该测试每个年代三部分中的正确条目数。三部分指:① 识别,正确识别名人数(随机概率水平为25%,如12.5/50);② 命名,无线索提示正确命名数;③ 确认,自主命名正确条目数(无线索提示)加上那些所提供的特定识别信息。

(三) 评价

该版本的著名人物面孔测验可用于检测AD患者的远期记忆损伤。

三、华山名人面孔命名测验

测试选取20个与受试者生活在同时代著名人物的照片,这些照片包括政治家、戏剧、电视及电影明星、主持人和运动员。其中政治家5位,戏剧、电视及电影明星10位,运动员4位,主持人1位。他们有的还健在,如胡锦涛、刘德华,而有些人已过世,如孙中山、毛泽东。所选照片均为黑白照片,且格式相同。在附表中注明每张照片的姓名,并匹配2个相同性别、同领域的著名人物

姓名。为了整体设计平衡，不同领域人物照片的位置都是随机的。

(一) 指导语

测试分为两个部分(命名、鉴别)。向受试者依次展示每张照片，让其说出著名人物姓名，记录结果；若受试者说错或无法说出该人物姓名，可进行线索提示，如“王楠”，提示“她是一名运动员”，记录结果；若仍无法回答，提供包括正确答案在内的3个选项，记录答案。只有受试者说出全名才算命名正确。

(二) 评分

评分包括命名和鉴别两部分。命名：有/无线索提示的正确命名数；鉴别：提供包括正确答案在内的3个选项后，回答正确数。

(三) 评价

用于AD与SD的鉴别诊断，有比较好的效度。目前尚在进一步验证中。

第四节　动作命名测验(ANT)

动作命名测验(action naming test，ANT)包括动态动作命名测验(dynamic action naming test)和静态动作命名测验(static action naming test)。

最近有研究对比了动态和静态动作命名测验，结果显示：两种类型的刺激在神经系统的处理上很大程度是重叠的，进一步验证了两种测验的意义可能无太大区别。Daniel等人为了进一步研究，剪辑了158个3～5 s的视频，要求受试者对每段视频给出最适当的动词描述，同时对每个受试者进行静态命名测验。研究发现：额叶盖和颞枕叶后外侧与动态动作命名测验的缺损有关(此区域损伤者无法完成动态动作命名测验)；动态动作命名测验和静态动作命名测验高度相关，在动态动作命名测验中失败者也总是在静态动作命名测验时失败。

本研究验证了动态和静态动作命名在神经系统有相当大的共同区域这一假说。

第五节　爱丁堡功能性交流能力检查法(EFCP)

在西方国家，经典的失语检查法包括波士顿诊断性失语检查(the Boston

diagnostic aphasia examination, BDAE)、西方失语成套测验(the western aphasia battery, WAB)和双语失语症检测法(the bilingual aphasia test, BAT)等。国内迄今发表的失语检查法都是参照国外传统的失语症检查法(主要是BDAE和WAB),结合国情,并考虑汉文化、语言特点及方言等因素,比较常用的有5个版本:中国科学院心理研究所胡超群等编制的"临床汉语言语测评方法"、北京医院王新德等编制的"汉语失语检查法(草案)"、北京大学第一医院高素荣编制的汉语失语检查法(aphasia battery of Chinese, ABC)、中国康复研究中心李胜利等编制的"汉语失语症标准检查法"和暨南大学附属第一医院陈卓铭等编制的基于计算机辅助的"汉语失语检查法"。以使用广泛的ABC为例,具体检查内容为:谈话、理解、复述、命名、阅读、书写、结构与视空间、运用和计算等。

这些检查法的优点是:以明确的"运动/感觉"、"表达/接受"两维模式取向为分析基础;能比较好地区分常见的失语类型,如运动性失语、经皮质运动性失语、感觉性失语、经皮质感觉性失语、混合性失语、传导性失语等;其失语分类和影像学证实的失语的解剖定位之间有比较高的符合率(67%~94%)。传统汉语失语检查法的局限性之一是把失语视为单纯的语言问题,而没有充分考虑其他认知功能(如记忆和执行功能)对语言功能的影响作用;失语治疗效果可能反映在"沟通能力指数(PICA)"和"功能沟通轮廓图(FCP)"上,而并没有反映在听、说、读、写、算为内容的失语检查法上。

语言交流的主要目的在于人际沟通。传统的语言功能标准化测试只是描述失语症患者交流能力的一部分,因为传统的评估为患者提供的是非真实的语言标准,是脱离人际环境孤立地进行评定的语言运用能力。Holland指出,患者语言的恢复往往取决于他能多好地适应残存语言功能,而不是在语言技能的再学习上取得了多大进步。大量证据表明失语症患者因神经系统损害而导致的语用学技能损害并不如语言技能损害重。Wilcox引用许多研究表明失语症患者保留了交流活动的起始和维持、上下文信息的编码和解码,以及对检查者的反馈做出反应的能力。

爱丁堡功能性交流能力检查法(Edinburgh functional communication profile, EFCP)是一种描述失语症患者交流有效性的观察法,1990年修订版,它的设计基于功能性交流就意味着"信息传递"的理念。在激发语言自然运用的环境下没有任务呈现的检测手段就是功能性的。EFCP在观察者之间的一致性信度、复测信度和效度良好。非言语行为的观察较之言语形式,无论在语言发展还是生理发生方面,都更原始,因此更能抵御脑部的损伤。尽管非语言

交流可能并不如认知模式那么受关注,但唯有它能提供交际水平上的应对策略。实际上,"聪明的"失语症患者可能运用的交流方式远多于 EFCP 中的描述,比如,手势、视线、点头、指示、画图。

一、EFCP 的发展目标和背景

大量证据表明失语症患者因神经系统损害而导致的语用学技能损害并不如语言技能损害重。Wilcox(1983)引用许多研究表明失语症患者保留了起始和保持交流活动、编码和解码上下文信息及对听者的反馈做出反应的能力。完成这些功能的策略可能包括:指出需要一些时间来回应;鼓励听者提供两者择一的选项;指出重复的需要;听者对不能理解的项目进行解释。这些策略能通过各种形式或其组合来实现。EFCP 的设计旨在通过精确地参考语言环境,对各种策略的形式及其实现程度进行结构观察和分析。

二、EFCP 的信度和效度

作为一种观察方法而不是标准化测试意味着某些个体差异是不可避免的。这当然也存在于日常面对面交流中。只要这些差异不是由于对观察项目的误解而致,那么作为观察者之间个体差异的反映是可以接受的,事实上还可能成为使人感兴趣的方面。尽管为了研究的目的可自该表得出量化的结果,但应避免用数字得分来表示,因为有迹象表明给定的分数会掩盖行为观察法所收集到的信息。表中的反应等级是按言语分类的,并不是与数字序列相关的连续体。2 个月的控制研究,表明观察者之间的一致性水平是可以接受的。大多数治疗师认为该检查表有助于评定患者在真实但又系统的方式下的总体交流能力,且在某些情况下是唯一可行的评估方法。

三、EFCP 与传统测试的关系

设计这一检查表并不是为了替代任何传统量表。当制定基于认知过程的任务导向疗法时——对应基于沟通过程的交流导向疗法,标准化量表依旧能提供最多信息。迄今为止的效度试验表明 EFCP 与传统测试之间存在负相关,正如预期所料。例如,患者可能在波士顿诊断性失语检测中得到令人沮丧的 z 得分——大部分亚项在−1 到−2 分之间。但是在 EFCP 测试中,却显示出有效地进行问候、确认、回应、要求等交流能力。这表明尽管存在残存语言交流困难,患者仍可胜任一系列的功能性交流。对其使用策略进行分析能够

为治疗提供指导。

四、EFCP的检查内容

(一) 交流形式

并非所有的老年患者都使用言语作为首选交流方法。患者在不同的情景下可能会使用以下5种交流形式：(S)言语；(G)姿势，包括形式化的手势；(F)面部表情；(V)不成字词的嘟囔之音；(W)书写。

如果可行的话，大部分患者都会运用言语来交流，但也补充以姿势和(或)面部表情的有效运用。如果一个策略包含多种形式，需要注明。

如果并不是一种形式补充另一种，而是运用了各种形式的组合——它们独立运用时都不能有效地交流，那么应当在检查表中注明，用括号把各项括起来。

某些形式可能并不适用于每位患者(比如，具有充分姿势和言语能力的患者无需运用书写)，这种情况下可留空而不需记录。

(二) 反应等级

患者的反应可以有很大的差异：一些病患的反应局限，另一些比较多样化。每个交流项目尽可能只记录一个最常发生的等级。如果对同一项目的反应变化很大，则有必要记录多次，并在注释栏中加以说明。反应分7个等级：无反应，答非所问，词不达意，刻板语言，达意，合格，详细。

1. 无反应(NR)　5种形式中任何一项都没有反应。需注意，某一形式的无反应要与广泛交流障碍相鉴别。如果仅仅是没有运用某种形式的交流，并没有不适当的行为(比如，某位患者没有用书写的形式来问候)，需要留空而不是记录为无反应。

2. 答非所问(Inap)　不相关的回答，比如，对患者打招呼：你好，Bloggs先生，你今天感觉怎样？他的回答是：

言语：谢谢你。

姿势：指向说话人。

表情：困惑的，皱眉。

发音：—

书写：写数字。

3. 词不达意(Inad)　不足以表达特定含义的局限回答。比如，问患者“你到达这里有没有困难？”，回答是：

言语：哦，呃……我……想……是的。

姿势：摆手(表示否认)。

表情：抬起眉毛。

发音：运用降调。

书写：潦草难辨。

4. 刻板(Ste) 如果患者出现刻板回答，需加以记录。这种回答的有效性差异很大，在一些情况下可能是达意的。比如问他“你今天怎样?”，回答是：

言语：很好。

姿势：握手。

表情：微笑。

发音：升调。

书写：写“好”。

每个形式的回答都达意。但是，当这些反应在交流过程中反复出现以造成刻板印象时，需记录为刻板。可通过观察各种交流形式及日后随访来确定这种形式能传递多少患者欲表达的信息。这些交流信息需在表格相应的可交流或不可交流项中记录下来。

5. 达意(Ad) 能传达特定含义但是交流比较局限。比如，问“你到达这里有什么困难吗?”，回答是：

言语：“哦，很好，没问题。”

姿势：指向门。

表情：—

发音：合适的音调。

书写：写“好”。

6. 合格(Qu) 较之达意的回答，还增加了一些内容，比如，对上述相同问题的回答：

言语：“啊，我想，还不坏。”

姿势：竖起大拇指。

表情：有目光接触，噘嘴表示有找词困难。

发音：运用语音停顿来表示某些不满意。

书写：写“我不得不使用我的拐杖”。

7. 详细(Elab) 提供完整的信息。比如，对上述问题的回答：

言语：“嗯，我想我尽力了，但是和以前当然没法比，走台阶困难。”

姿势：使用合适的手势，摇头。

表情：有目光接触，运用表情表示不满意（皱眉，噘嘴）。

发音：运用语音停顿表示正在思考。

书写：写“没有运用。”

8. 无证据（NE） 在观察过程中没有充分的证据来评价患者某一形式的反应。这种情况下应记录为无证据而不是无反应，以便在日后可以再次评定。

除了以上所述的各种等级，患者在各个形式上的反应可能是不一致、相互组合的。

9. 不一致的反应 通常，患者的反应包括了有效性不同的各种方式。要注意避免这种倾向——只记录能成功交流的形式。比如，招呼患者：“早上好，Bloggs 先生。”回答是：

言语：“呃……电话（Inap）。”

姿势：握手（Ad）。

表情：目光接触，微笑（Ad）。

发音：升调（Inad）。

书写：写“没有运用。”

与此相对的是，详尽的自我中心的言语加上不适当的非语言形式。比如，问患者“你今天感觉如何？”，回答：

言语：“当然，我目前的健康状况比起几个月前好多了。但问题是我真的很需要四处走动，现在这种情况却不允许我这么做（Elab）。”

姿势：摇头，挥动手臂（Inap）。

表情：皱眉，面无表情，目光闪躲（Inap）。

发音：过度的音量（Inap）。

书写：写“没有运用（留空）。”

10. 组合 患者的反应可以由各种形式组合，任一形式在独立运用时都不能有效交流但相结合后就组成了一个成功策略。在检查表中用括号把 2 个或更多的形式括起来。比如，提到患者的丈夫时，问患者“他是个好厨师，是吗？”，回答是：

言语：“嗯……呃……锅……啊……锅。”
姿势：做出持锅的姿势，当尝试记起一些词摇头。
表情：目光、嘴唇的动作表示不满意。 } Qu
发音：降调。
书写：写“没有运用。”

每个形式独立地看，几乎连达意都够不上，但是当组合起来后就能传达较多意义。在这个例子中，可以将言语、姿势、表情、语音几项括在一起，评定为合格。

(三) 交流内容

该检查表分为 6 个交流方面：问候、确认、回应、要求、提议和口头解决问题。每个部分都包含了交流能力的不同等级，除了第六项——只运用于那些具备有效交流能力的患者。

1. 问候　有人进门时，患者的觉察反应。

他是否望向别处，没有回应？—yes→检查他是否能听到或看到你。—no→记录为 NR，再次问候，他是否有反应？

患者是否虽然没有说话，但是注视你或者对你微笑挥手等？—yes→在得分纸以下各项中(GFV)记录为 Ad 或 Ste。

患者是否对常用问候语，比如，"你好"、"早上好"做出简短回答？—yes→在相应项记录为 Ad：SGFVW。

患者是否用一些短语发表对特定地点或人物的评论，比如，"今天这里很冷"，—yes→在相应项记录为 Qu 或 Elab：SFGVW。

"我喜欢你的发型"，"你感觉好些了吗？"确认他的反应不符合前述各项。

2. 确认　患者做出反应——确认其他人说的话语已被理解。这种确认能够通过多种形式表示，但是必须保证其可靠性以作为进一步交流的基础。

配偶或护士报告的反应必须要附以具体事例做参考，以防止过分乐观的结论。将这些信息以"来自报告"记录在注释栏中。

当你对患者说话，要求他做出确认的时候，他是否有反应？—no→重复话语后他是否有反应？—no→记录为 NR。

虽然保持沉默，但是点头或摇头，或运用脸部表情或语音来表达意思。—yes→在相应项中记录为 Ad：GFV。

他是否回答"是/否"问题？—yes→在得分纸相应项记录为 Ad：SGW。

重复 2～3 次以引出更多回答，记录下反应的形式和等级。

3. 回应　患者在以下各种情形下的回应：① 非语言形式的问题。② 闭合式问题，比如：是/非问题，正/误判断。③ 开放式问题，比如：特殊疑问，需患者解释、表示认可或同意的问题。④ 描述，比如，描述他的家/他玩的游戏/如何煮汤。

患者是否对有意义的姿势做出回应？—no→对非语言形式的要求：NR。

患者是否对是非问题作出回应？—no→记录下对非语言形式要求的回答及其形式等级。闭合式问题：NR。

患者是否回答需表达个人意见的问题？—no→记录下对闭合式问题的反应形式及其等级，开放式问题：NR。

患者是否能描述病房、医院、他的家或一幅图画？—no→记录下开放式问题的反应形式及其等级，描述性问题：NR。

4. *要求* 患者提出要求的能力怎样？通过观察患者或(和)配偶及护理人员的会谈后记录。

患者吸引他人注意和提出要求的能力都要考虑在内：① 引起注意：作为提出要求的准备步骤，患者是否尝试着引起别人的注意。对于只能通过非语言形式来提出要求的患者，这一步尤其必要。② 索要物品，要求做某事：患者如何表示他想要做某事，比如，他想去洗手间，他想睡觉，他饿了，渴了，想关掉电视机，等等。③ 要求信息：患者如何表示他想得到信息，比如，关于来访者、钱、管理等方面的信息，或者要求提问者重复一下问题。

患者是否有想要交流的表示？—no→在“吸引注意”一栏记录为 NR。

患者提出的要求可被理解吗？—no→在“吸引注意”一栏记录为 Inad 或 Ste(不可交流)。

他是否尝试吸引？如果有人为引起注意，记录为 Ade，并记下其形式。如是否有特定人员的注意？—no→走近他，他是否有特定交流？—no→如果反复出现则记录为 Ste。

他是否把他人的注意引向他要求做某事或要求得到信息？—no→在“吸引注意”一栏中记录为 Qu，并记下其形式。

记录下“索要物体，要求做某事，要求信息”的形式及其等级。

5. *提议* 这一项包括患者引出话题与他人交流。提议能通过以下形式表达：① 传递物体，比如，他如何表达希望传递给听者一个物品，比如，一个空茶杯。② 引出话题。

他是否有自发传递物体的行为，记在“物品传递”一栏中，或有传递物体的愿望？—yes→记录下反应的形式及等级。

他是否引出话题？—yes→在引出话题一栏中记录下反应形式及等级。

他是否持续地以自我为中心地谈话而不顾及他人。—yes→在引出话题栏中记录为 Ste(可交流)，并在注释栏中记录。

6. *口头解决问题* 这个部分只用于具备有效交流能力的患者，要求患者

在谈话中能运用一些口头形式来解释问题。比如：① 日常生活情景的问题："你今天是怎么到达这里的？"② 什么样的帮助对残疾人是有用的？你能想出一些应对这些问题的办法吗？

（四）检查表和得分纸

EFCP 是为了提供非正式情景下或功能性交流而设计的检查表。它可以与其他标准化失语成套测验联合运用，互为补充，以提供受损语言能力的精确信息。

EFCP 作为标准测验的补充，提供了更多关于 AR 的精神状态和交流能力的信息，尽管针对语言功能的测试项目相对较少。它描述了传递信息的形式、传递的程度，以及这些形式的交互程度。实际上，它检查了那些语言能力受损而交流能力尚存的患者。

五、EFCP 的评分表或样例

EFCP 的评分表或样例见表 5-5-1。

表 5-5-1　EFCP 的评分表或样例

形　式		不可交流				可交流				注释 延迟 形式 组合	
		NR	Inap	Inad	Ste	Ste	Ad	Qu	Elab		NE
问候	S										
	G										
	F										
	V										
	W										
是非问题确认	S										
	G										
	F										
	V										
	W										
非语言形式的要求	S										
	G										
	F										
	V										
	W										

（续表）

形 式	不可交流				可交流				注释	
	NR	Inap	Inad	Ste	Ste	Ad	Qu	Elab	延迟 形式 组合	NE
闭合式问题	S									
	G									
	F									
	V									
	W									
开放性问题	S									
	G									
	F									
	V									
	W									
描述	S									
	G									
	F									
	V									
	W									
引起注意	S									
	G									
	F									
	V									
	W									
索要物品或要求做某事	S									
	G									
	F									
	V									
	W									
信息	S									
	G									
	F									
	V									
	W									

（续表）

形式		不可交流				可交流				注释	
		NR	Inap	Inad	Ste	Ste	Ad	Qu	Elab	延迟 形式 组合	NE
物体传递	S										
	G										
	F										
	V										
	W										
引出话题	S										
	G										
	F										
	V										
	W										
口头解决问题	S										
	G										
	F										
	V										
	W										

注：表中的缩略词意义。① 形式：言语(S)、姿势(G)、表情(F)、发音(V)、书写(W)。② 反应：无反应(NR)、答非所问(Inap)、词不达意(Inad)、刻板(Ste)、达意(Ad)、合格(Qu)、详细(Elab)、无证据(NE)。

（郭起浩　马　洁）

第六章

注意的评估

注意是指人的心理活动对一定对象的指向和集中。指向和集中是注意的基本特点。注意力就是把自己的感知和思维等心理活动指向和集中于某一事物的能力。人类大脑的容量是有限的，有效的信息处理必须对信息进行选择。通俗地说，“注意”就是信息流进大脑的那扇门。但是，各种有关注意的理论模型并不简单，而是相当复杂的。

注意的评估工具包括韦氏记忆测验的注意分测验[心智、数字广度测验(DST)、视觉记忆广度测验]、简易注意测验(brief test of attention)、同步听觉连续加法测验(paced auditory serial addition test，PASAT)、持续操作测验(CPT)、数字划销测验(如 Ruff 2&7 选择注意测验)、字母划销测验、符号数字模式测验(symbol digit modalities test，SDMT)、日常注意测验(test of everyday attention，TEA)、注意变化测验(test of variables of attention，TOVA)和连线测验(TMT)。本章介绍 DST、SDMT 和 PASAT，TMT 将在执行功能部分介绍。

第一节　数字广度测验(DST)

数字广度测验(digit span test，DST)是在测试者读出一系列数字后，检测受试者以正确顺序顺背、倒背该条目的能力。DST 是 WAIS 各种版本与 WMS 各种版本的一部分，这是唯一一个既出现在韦氏智力测验中，又出现在韦氏记忆测验中的分测验，也许韦氏(Wechsler)认为 DST 既能够反映智力也能够反映记忆能力。同时，它也是一种常见的测试工作记忆的手段，工作记忆是指保持多个信息处于在线状态，使人们可以在短时间(3 min)以内用这些信

息做一些事情，这一信息加工是基于监测和更新过程。

一、指导语

测试者说出一些数目，受试者仔细听，当测试者说完时，受试者就跟着照样背出来(以每秒 1 个数字的速度匀速读出)(表 6-1-1)。

表 6-1-1 数字广度测验(DST)

顺背		倒背	
3	5-8-2	2	2-4
	6-9-4		5-8
4	6-4-3-9	3	6-2-9
	7-2-8-6		4-1-5
5	4-2-7-3-1	4	3-2-7-9
	7-5-8-3-6		4-9-6-8
6	6-1-9-4-7-3	5	1-5-2-8-6
	3-9-2-4-8-7		6-1-8-4-3
7	5-9-7-1-4-2-8	6	5-3-9-4-1-8
	4-1-7-9-3-8-6		7-2-4-8-5-6
8	5-8-1-9-2-6-4-7	7	8-1-2-9-3-6-5
	3-8-2-9-5-1-7-4		4-7-3-9-1-2-8
9	2-7-5-8-6-2-5-8-4	8	9-4-3-7-6-2-5-8
	7-1-3-9-4-2-5-6-8		7-2-8-1-9-6-5-3
10	5-2-7-4-9-1-3-7-4-6	9	6-3-1-9-4-3-6-5-8
	4-7-2-5-9-1-6-2-5-3		9-4-1-5-3-8-5-7-2
11	4-1-6-3-8-2-4-6-3-5-9	10	6-4-5-2-6-7-9-3-8-6
	3-6-1-4-9-7-5-1-4-2-7		5-1-6-2-7-4-3-8-5-9
12	7-4-9-6-1-3-5-9-6-8-2-5		
	6-9-4-7-1-9-7-4-2-5-9-2		

1. 数字广度测验——顺背　受试者重复一系列随机数字，速度为一秒钟读出 1 个数字，一般是从 3 个数字的条目开始，不断增加长度继续测试直至 12 个数字的条目或连续 2 次回答错误为止。受试者被要求以正确顺序重复所有数字，如果受试者知道某数字的位置但忘记该数字是几，就说“空”。任何一个项目中的两个测试全部没有答对时即终止。

从哪个数字开始，在临床医生的实践中，可以灵活处理，为了节约时间，可以从 5 个数字串开始，通过则往更长数字，失败则往更短数字。当然，如果是

多中心的科研评估,必须按照事先的操作规程执行,不允许这样变通。

2. *数字广度测验——倒背* 该测试要求受试者以倒序方式重复测试者所说数字。每组数字由数字1到9随机组成。按如下内容指导患者:"现在我将说出更多的数字,不过这次我停下来时,我要你由后向前说出这些数字。例如,如果我说2-8-3,你该如何说?"如果回答正确,则说:"对。现在听这些数字,请记住你要由后向前说出这些数字。"如果回答错误,则说:"不是这样,我说2-8-3,那么由后向前说时,你应该说3-8-2。现在试试这些数字。请记住你要由后向前说出这些数字。准备好了吗? 6-2-9。"

二、评分

(1) 每回答正确一串数字计1分,总分:顺背____+倒背____=____这是最常用的分析方法;也可以分析倒背与顺背的比例。

(2) WAIS 的评分手册有粗分转化为标准分的方法。然而,正常人的 DST 的顺背得分范围比较窄,89%为5~8分(即6±1),标准分还没有粗分直接可靠。

(3) 分析指标还包括最长的一次通过的数字串、最长的第二次测试才通过的数字串,两者的比例。

(4) 错误类型:① 心理追踪困难:在数字串的中间搞混数字次序;② 数字替代:如3-5-6代替3-5-9;③ 持续错误;④ 数字减少,如4-8-2-9代替4-8-2-9-5。

三、评价

因为随机数字容易生成,很容易做成 DST 的不同版本,适合药物临床试验的前后比较或者需要复测的检查。DST 数字阅读时应该均速,否则,容易产生"组块"策略记忆(如3-2-6-8,记为32-68)。

Ardila(1989)提出顺背5分是正常的下限,4分是边缘状态,3分有肯定的损害。由于中文数字的发音与英文不同,中文数字是单音节,相对来说,相同长度的数字音节比较少,这可能有助于中国受试者的 DST 得分提高,相应的中国人的 DST 的划界分至少比年龄、教育匹配的西方人高1分。

患者 DST 失败的影响因素可能是分心、配合不佳、不理解测试要求,比如,有些患者2次复述失败,以为有第3次复述,而如果进行第3次复述,该项目复述可能正确完成。如果第3次复述能够完成,那么,至少能够完成延长

1～2个数字的项目。如果患者的倒背优于顺背，显然是患者的主观努力不够。尽管DST的低分可能与情绪（如焦虑、应激）有关，有研究发现不同的焦虑水平的正常人的DST得分并没有显著差异。

Waters等人研究表明年龄对DST的倒背有明显影响，对70岁以下正常人的顺背几乎没有影响，70岁以上的影响也非常小。Craik（1990）报道84～100岁健康老人的顺背可以达到5.7±1.0，范围4～8分。而教育程度对顺背有决定性影响。数字广度测验-倒背与画钟实验有高度相关性。

DST检测即刻记忆和注意力。在AD的神经退行性病变进程的早、中期，即刻记忆能力没有显著下降，因此，有的研究者将DST作为AD早期其他测验的阴性对照。但是，也有研究者认为AD早期的DST顺背数字长度已经出现引人注目的减少。

总之，DST尽管操作简便、耗时少，但它的局限性也是明显的：它对于认知障碍的早期诊断不够敏感，表现受文化背景的影响。

四、其他版本

DST有许多相似的版本，如点数广度（point digit span，Smith，1975）、字母广度（letter span）、Knox立方体测验（Knox cube test，Arthur，1947）、Corsi积木敲击测验（Corsi block-tapping test，Milner，1971）、句子复述、次序操作系列（sequential operation series）等。

数字次序测验（adaptive digit ordering test，DOT-A）是由Cooper等1991年编制的一种新型数字次序测验，用于分析数字广度。Katja Werheid等（2002）采用DOT（表6-1-2）对帕金森和额叶损伤患者进行数字顺序测试，发现这些患者常常表现为执行功能障碍中的工作记忆损害，表现为DOT-A得分低，而DST顺背得分则无明显变化；DOT-A与DST倒背得分在同一水平。DOT-A对帕金森患者有独特敏感性。

表6-1-2　数字次序测验（DOT）

Digit Ordering-A		测试1	1/0	测试2	1/0	得分	0/2
条目	广度长度						
1	3	3 7 2 (2 3 7)		6 1 6 (1 6 6)			
2	4	8 4 7 3 (3 4 7 8)		7 2 7 6 (2 6 7 7)			

（续表）

Digit Ordering-A		测试 1	1/0	测试 2	1/0	得分	0/2
条目	广度长度						
3	5	2 6 5 3 1 (1 2 3 5 6)		5 1 8 5 2 (1 2 5 5 8)			
4	6	4 8 3 9 2 7 (2 3 4 7 8 9)		3 0 3 7 2 5 (0 2 3 3 5 7)			
5	7	7 4 3 1 8 2 6 (1 2 3 4 6 7 8)		9 6 5 2 8 6 3 (2 3 5 6 6 8 9)			
6	8	3 7 6 1 8 2 5 0 (0 1 2 3 5 6 7 8)		2 6 3 2 1 8 7 5 (1 2 2 3 5 6 7 8)			
备注：				得分(12 分)： 得分百分比： 最大跨度：			

Digit Ordering-B		测试 1	1/0	测试 2	1/0	得分	0/2
条目	广度长度						
1	3	4 8 3 (3 4 8)		7 2 7 (2 7 7)			
2	4	9 5 8 4 (4 5 8 9)		6 1 6 5 (1 5 6 6)			
3	5	3 7 6 4 2 (2 3 4 6 7)		6 2 9 6 3 (2 3 6 6 9)			
4	6	3 7 2 8 1 6 (1 2 3 6 7 8)		4 1 4 8 3 6 (1 3 4 4 6 8)			
5	7	6 3 2 0 7 1 5 (0 1 2 3 5 6 7)		8 5 4 1 7 5 2 (1 2 4 5 5 7 8)			
6	8	4 8 7 2 9 3 6 1 (0 1 2 3 5 6 7 8)		1 5 2 1 0 7 6 4 (0 1 1 2 4 5 6 7)			
备注：				得分(12 分)： 得分百分比： 最大跨度：			

说明：在以上测试中，测试者将随机读出一些数字，请按从小到大的顺序重复出该组数字，如果数字出现两次，请重复两次。例如：5－2－8－2，答案：2－2－5－8。以相同音调、每秒钟 1 个数字的速度读出每条，在读每条的最后 1 个数字时音调降低，任一条目的 2 条均回答错误，测试终止。

DOT-A 是由逐渐增加长度的 6 个条目组成(3～8 个数字)，每个条目有 2 个数字串，其中第二串有重复数字。测试者以每秒 1 个数字的速度读出条目后，要求受试者立即以从小到大的顺序重复这些数字，计分方式同 DST。如果

同一长度的 2 条测试题均错误，则测试终止，因此，本测试与 DST 唯一不同的是数字的顺序。如果最后回答的条目中只有 1 条回答正确，减去 0.5 分，比如：如果 5 个数字的条目中只有 1 条回答正确，最大广度长度计分为 4.5 分。

Cooper 等人认为 DOT-A 是一种有前途的诊断工具，它的组成、与 DST 的直接可对比性及对患者的敏感性使其非常适合临床实践。

第二节 符号数字模式测验(SDMT)

符号数字模式测验(symbol digit modalities test，SDMT)是 Aaron Smith 1973 年发表、1982 年修订的，用来评估注意分割、视觉扫描、跟踪和运动速度。测试中的编码键(code key)包含 9 个不同的抽象符号，每个符号与一个数字相对应。要求受试者看过编码键后，以最快的速度写下对应于每个符号的数字。测验可用书写或口头方式进行。书写形式也可用于团体测验。在文字版本中，患者应根据页面上方编码键中提供的信息，填入与符号相对应的数字。在口语版本中，测试者记录患者说出的数字。当两个版本都施行时，建议先施行文字版本。

一、指导语

将测试表(图 6-2-1)放在患者面前，然后说："请看这些方格(用手指着表的上方的图解)，可以看见上面一行的每一个方格里都有一个符号，符号下面一行的方格(用手指着)，有对应的数字。根据这个配对关系，请您在每一个空的方格内填上数字，数字必须与上面的符号配对。例如，当您看了第一个符号标记后，您将会发现 1 可以写到第一个空格中。如果填错了，请不要涂抹，直接将正确的答案填在错误答案上就行。现在请你练习一下，将余下的空格填完，填到双线处停止。"在练习时，为患者指出所有的错误，并加以纠正。注意，如果患者不按顺序填空，应提醒他或她不要跳着填。继续测验，告诉患者："现在，当我说开始时，请您立即按刚才练习过的方法开始填写数字，越快越好，一直到我叫停为止。不要跳格，填得越快越好。准备好了吗？开始。"准确计数 90 s，然后说："停！"

对于不能书写的受试者，可以要求口头说出配对的数字，由测试者记录。

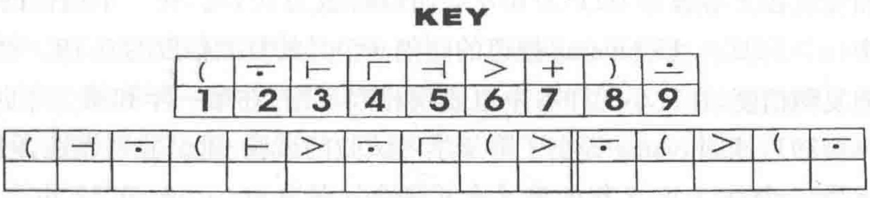


图 6-2-1　符号数字模式测验之模板

Feinstein 发展了 SDMT 的计算机版，Uchiyama(1994)要求在完成 90 s 配对后附带回忆与每个符号相配的数字。在 Uchiyama 的附加回忆版中，在标准流程的基础上，会给予受试者一张新的纸，上面有一行 15 个符号，包括译码的 9 个符号。Miller(2005)提出符号的次序可以根据 SDMT 最后一行组成。

二、评分

90 s 内正确填写的个数为最后得分，不包括在练习时填的数字，填写错误的数字个数也应记录，最高分为 110 分。

三、评价

1. 影响因素

(1) 年龄：SDMT 分数不管文字版还是口语版都随着年龄的增长而降低，或许反映了运动反应速度和信息处理速度、视觉查找和记忆的改变。

(2) 性别：Smith(1991)报道性别会影响表现，通常女孩的表现较男孩更好。这在非西方国家孩童中也得到验证。总体来说，在 8～13 岁的儿童中，无论男孩还是女孩，口语皆比文字形式表现好。然而两性之间的差异随着年龄增长而逐渐消失，尤其是 14～17 岁的青少年。成人中，虽然有些报道提示女性较男性表现好，但并非所有研究都发现两性间的差异。在一项大型的社区研究中(n>7 000)，女性 SDMT 的表现优于男性；当对于教育、健康状态等影响因素做出校正后，这个差别更突出，也就是较好的教育及健康因素不能解释女性表现较好的情况。

(3) 教育：受过良好教育的个体(13 年以上)较低教育程度个体(12 年以下)的表现好。根据这些不同，SDMT 需要提供对于年龄和教育程度(12 年以下对比 13 年以上)的不同标准。

2. 信度与效度　SDMT 的复测信度是相当不错的。Smith(1991)报道对 80 位正常成人进行 2 次文字版和口语版的测试，中间平均间隔 29 天。复测信

度的相关性在文字版 SDMT 为 0.80，而口语版为 0.76。在一个男性的大型研究中($n>1\,000$)，Uchiyama 报道的间隔 6 个月的复测信度为 0.79。对更长间隔的复测信度，在较小的样本中也表现相当不错(在第一年和第二年时 $r=0.72, n=39$)。Uchiyama 提出了附加学习试验的试验-重试验可靠性，但并不是非常强(6 个月、1 年、2 年的重试验可靠性分别为 $r=0.46$、0.52 和 0.71)。测验手册中说明在复测时会增加 4 分左右。Hinton-Bayre(1997)使用替代版本，发现在重测验中正确数目会有 2%的增长。但是 Uchiyama(1994)的超过 2 年的时间间隔复测文字版本并没有发现任何显著的练习效应。Hinton-Bayre(1999)发现练习效应的大小取决于复测的间隔时间。

文字版与口语版存在高相关性，但并不完全可转换使用。在正常成人中，文字与口语版本的相关性达到 0.78 以上。相同的，Ponsford 和 Kinsella(1992)在颅外伤的患者中报道了 0.88 的相关性。但两个版本的平均得分有差异。Yeudall 在 15~40 岁的 225 名志愿者中发现口语版的得分较高(约高 11 分)。

SDMT 用于评估注意的扫描和轨迹，其作用与数字划销测验、连线测验、WAIS 的数字符号分测验和反应时间测验相似。

3. SDMT 与 DSST 比较　SDMT 的测验材料和操作过程与 WAIS 的数字符号分测验(digit symbol-coding subtest)相似，前者在符号下写数字，后者在数字下写符号。符号都是 9 个无意义几何图形。WAIS-DS 的回忆有两种：偶然学习(incidental learning)要求受试者就数字回忆匹配的符号；自由回忆(free recall)要求写下尽可能多的符号。

SDMT 与 WAIS-DS 的相关性在不同人群中介于 0.62 到 0.78 并达最高的 0.91。虽然他们有非常高的相关性，但 SDMT 的原始分数一般较 WAIS-DS 低，这可能是由于 SDMT 难度的提高。由于 SDMT 需书写或表达的是已熟悉的数字，它的书写反应会较书写不熟悉符号的 WAIS-DS 来得容易许多。但从注意力的角度，WAIS-DS 比 SDMT 来得容易，因为，在 WAIS-DS 中，空间位置的线索都包含在编码键里，因为数字根据大小排列在页面上，容易找到。在 SDMT 中，符号的出现是随机的，编码键中没有任何提示位置的线索，符号之间没有内在联系。

4. 临床应用　SDMT 已经广泛用于各种神经疾病患者，如头部外伤、癫痫、帕金森病、脑血管意外，也用于药物依赖、衰老、内科疾病和中毒性疾病等的脑损害程度的评估。

SDMT 对成人和儿童的颅脑损伤格外敏感，尤其因为它的敏感性及在研

究中的高使用频率，它已成为在脑外伤(TBI)、多发性硬化(MS)、亨廷顿病(HD)和脑震荡中最常用的标准测试之一。表现的下降也与其他一些情况相关，包括癫痫、卒中、有机溶剂暴露、帕金森病；年龄、运动、脂肪摄入、整体健康，还有，事件相关脑电位的P300；物质滥用、精神分裂症、睡眠呼吸障碍和酗酒者的认知功能障碍。它对患有进展性脑肿瘤的成人也有识别作用，也被用于在机组人员中评估机油对认知功能影响的神经毒性研究，以及用于癫痫研究中评估抗癫痫药对认知功能的损害。Kinsella发现Stroop色词测验、同步听觉连续加法测验、反应时间测验和SDMT都是脑外伤患者信息加工速度的良好的预测因子，但SDMT对加工速度减慢最敏感。SDMT对运动员中脑震荡的影响敏感性很高，包括随时间的改变及不同严重程度中的康复率。在运动员中，它是决定是否回到赛场的标准指标之一。SDMT也被用来鉴别MS的类型。在一项研究中，作为MS患者一套全面评估组的一部分，它是筛查MS患者中神经心理障碍敏感性最高的测试。Pfeffer应用MMSE、瑞文推理测验和SDMT等鉴别痴呆和认知正常老人，发现SDMT是最佳鉴别指标。

SDMT具有容易执行、快速、可靠性高的优点。它的口语版本优势在于对手功能障碍的个体也能进行测试，而它的文字版本适用于口语障碍的个体。用于团体测验，它快速又经济，并且可用于筛查儿童或成人神经认知障碍，或是用于特殊情况，就像是获取运动员赛季前的基线表现。最重要的是，它是神经心理学中最敏感的测试，有大量的研究支持它识别认知障碍、功能改变、疾病进展和不同临床患者群残疾状况的能力。它对于认知障碍的敏感性也使它成为许多临床疾病的标准化成套测试之一。

就像Smith提出的，文字与口语的替代试验需要结合许多复杂的认知能力，比如视觉功能、运动功能、口语语言能力和注意功能。然而Bate提出由于SDMT的多因素复杂性，因此很难鉴别是特定的注意力不足还是信息处理速度减慢。这可能也是SDMT相较于其他具有更细分能力范围测验的优势，因而它在不同人群中都有很好的敏感性。

（郭起浩）

第三节 同步听觉连续加法测验(PASAT)

同步听觉连续加法测验(paced auditory serial addition test，PASAT)是

通过一系列加法任务评估受试者的工作记忆、注意分配和信息处理速度。该测试的适用年龄很广。Gronwall 版 PASAT 的适用年龄范围为 16～74 岁，Levin 版为 20～68 岁，计算机版适用年龄由其原版决定。儿童版同步听觉连续加法测验(CHIPASAT)的年龄范围为 8～14.5 岁。

一、PASAT 版本

PASAT 最初被 Gronwall 等用于评估头部外伤患者的信息处理速度。该测试是根据 Sampson 1956 年版发展来的。虽然 1956 年版还包括视觉呈现项目，但听觉呈现项目更常用于神经心理评估。

各版本的测试，都是逐渐呈现一系列随机数字(1～9)，受试者需要在数字出现后立即将相邻的两个数字相加，如第一个数字与第二个数字相加，第二个与第三个相加，第三个与第四个相加。举例来说，第一个数字为 1，接着出现第二个数字 9，受试者需要立即回答它们的和是 10，之后出现第三个数字 4，受试者需要回答 13(答案为第三个数字和上一个数字之和，而不是第三个数字与之前一二个数字的和相加)，然后按这样的规律继续直到该组测试结束。然后继续相同的测试，但数字出现的时间间隔(interstimulus interval，ISI)会缩短，共 4 组。PASAT 通过加快刺激时间、缩短反应时间来增加难度。

PASAT 要求受试者对每个数字集中注意力，通过个位数的计算运用工作记忆，适应逐渐增加的难度，减少反应时间。PASAT 是评估注意分配、注意保持和工作记忆的有效测验。迅速完成测试任务也需要信息处理速度。当然，该测试同样需要基本的计算能力，但不同于以前评估初级信息处理速度的概念化测试，PASAT 被认为是一项高要求的多因素测试。

现在还有很多不同版本的 PASAT，主要区别在于刺激的感觉模式(听觉或视觉)和刺激的形式(录音带或电脑)。测试的数字个数和实际刺激间隔时间(ISIs)也与标准版不同。大多数版本是模仿 Gronwall 版(每组测试 61 个数字)或 Levin 版(每组测试 50 个数字)。简短版(录音带版)PASAT 只测试 1 次(PASAT - 50)或 2 次(PASTA - 100)。儿童版数字相加的和不超过 10。电脑版通过听觉或视觉呈现数字，故记录受试者的回答方式也不同(人工记录 vs. 电脑记录)。

与中国有许多方言一样，在讲英语的不同地方，人们的发音也有口音的差异，所以听觉版 PASAT 的英语口音存在地域区别，如 Gronwall 版的口音为

新英格兰口音，Levin 版的口音为美国口音，其他还有加拿大、英国、法裔加拿大口音。

也有其他版本已经被用于 fMRI 研究或成为 MS 研究的核心评估工具。基于计算机辅助的版本称为调整版 PASAT(adjusting - PASAT, Tombaugh, 1999)，可以根据受试者的表现调整刺激间隔时间，同时减轻受试者的紧张情绪。

Deary 于 2001 年编制的非数字的 PASAT，称为同步听觉连续对立测验(paced auditory serial opposites task, PASOT)，受试者听录音带播放的一系列单词，受试者每次听到一个单词，要求他(她)说出前一个单词的反义词，例如，大—热—(受试者说小)—高—(受试者说冷)—快—(受试者说低)—上—(受试者说慢)，以此类推，共 10 个常用单词，每隔 4 s 播放一个。移除了计算步骤，编制者认为消除了焦虑之源，能够更好地反映注意力。

二、测试方法

以 Gronwall 版为例，指导语在录音带里。播放时音量调节到受试者能够清晰听到。对某些病情严重的受试者，测试者需重复或解释指导语。记录纸见表 6 - 3 - 1。

测试先从练习开始，再进行第一组测试(ISI 2.4 s)。在之后的每组测试前，测试者需告知受试者下一组测试各数字间的间隔时间比上一组更短。两组测试间隔时间至少 60 s。很多受试者被发现即使是较慢的测试(2.4 s 和 2.0 s)完成都很困难，故只有在 2 组较慢的测试中表现良好的受试者才进行 2 组较快的测试(1.6 s 和 1.2 s)，如 2.4 s 间隔时间测试正确数超过 40 个或 2.0 s 间隔时间测试正确数为 20 个以上。

其他版本如 Levin 编写的普通版、Diehr 等编写的简短版(PASAT - 50、PASAT - 100、PASAT - 200)、CHIPASAT、电脑版 PASAT、调整版 PSAT(Tombaugh, 1999)，它们的指导语与测试表大同小异，由于 PASAT 几乎全部版本都是免费的，有兴趣的读者可以很容易地在网络上检索到具体内容。需要注意的是，当发现受试者在使用“组块”策略(选择性回答以规避工作记忆)时，需要停止测试并重复指导语。

完成不同版本需要的测试时间不同。Gronwall 版的所有 4 组测试共耗时 15～20 min。Levin 版的时间更简短，所有 4 组测试耗时约 11 min。简短版(PASAT - 200、PASAT - 100、PASAT - 50)只耗时 6～8 min。

表 6-3-1　Gronwall 版 PASAT 的记录纸

	2.4 s	2.0 s	1.6 s	1.2 s		2.4 s	2.0 s	1.6 s	1.2 s		2.4 s	2.0 s	1.6 s	1.2 s
7(9)					8(12)					5(13)				
5(12)					7(15)					4(9)				
1(6)					1(8)					8(12)				
4(5)					6(7)					2(10)				
9(13)					3(9)					1(3)				
6(15)					5(8)					7(8)				
5(11)					9(14)					5(12)				
3(8)					2(11)					9(14)				
8(11)					7(9)					1(10)				
4(12)					5(12)					3(4)				
3(7)					3(6)					6(9)				
2(5)					4(7)					2(8)				
6(8)					7(11)					9(11)				
9(15)					1(8)					7(16)				
3(12)					5(6)					8(15)				
4(7)					8(13)					2(10)				
5(9)					3(11)					4(6)				
8(13)					4(7)					7(11)				
6(14)					6(10)					6(13)				
4(10)					8(14)					3(9)				

三、评分方法

1. 总正确数 最常用的评价 PASAT 表现的指标。需要注意的是，使用总错误数（不包括遗漏数）而不是正确数可能会令人误解，因为随着测试难度的增加，受试者只能回答更少项目，错误数也会逐渐减少。

对 Gronwall 版评分时，测试者需记录每一组测试的正确数。每组测试最大的正确数为 60（总正确数最大为 240）。测试者需要计算所有测试错误答案的比例（包括遗漏数）。错误答案比例需小于 10%，若超过 20%，则显示 PASAT 可能无效。实际错误答案数可能比遗漏数少。举例来说，Wingenfeld 等（1999）的研究中只有一个受试者符合错误数超过 20%的排除标准。

大多数 Levin 版或每组测试 50 项目的版本使用每组测试的正确数作为主要结果指标。使用 Diehr 等（2003）编写的简短版时，需要计算所有组测试的总正确数（最大正确数为 196），然后根据基本资料用总正确数换算量表分。

评分 CHIPASAT 时，需计算总正确数，也可以计算每组测试的正确数。有研究觉得每组测试的正确数更能反映信息处理能力。

2. 每个正确回答所需时间 将总正确数转换为每个正确回答所需时间来反映受试者的信息处理速度可能不正确。这个方法为测试的持续时间除以测试的正确数。但问题是测试结果中实际上并没有提供受试者回答的速度。如果需要评估信息处理速度，则计算受试者实际回答时间更合适，而只有电脑版可以记录受试者的实际回答时间。

3. 二数分（dyad scoring）/组块分（chunking scoring） 附加分也有高灵敏度。其中之一为二数分。当受试者连续回答正确两个项目即二数分得 1 分。二数分可以测定受试者是否能尽量完成每一个项目而不是采用“组块”策略来提高分数（如回答间隔的项目而不是回答连续的项目）。只回答间隔的项目会显著降低测试任务的工作记忆成分，延长每个正确回答所需时间，错误估计信息处理速度。二数分百分比可通过二数分除以总正确数再乘以 100%计算得到。

组块分即为遗漏项后第一个正确回答数。随着测试难度的增加，临床患者和正常对照组都倾向于使用组块策略。需要注意的是，测试者需要将受试者的回答分为不正确、正确、遗漏，以此获得二数分和组块分。

由此可见，二数分反映了受试者遵从测试指令要求，而组块分反映了随着测试难度的增加，受试者采用的补偿策略。二数分百分比并不反映受试者表

现的准确性，而反映了受试者跟上指令的能力。例如，被试总正确数低，二数分高，说明受试者跟上指令的能力强，相比通过组块策略强行跟上指令的受试者具有更好的信息处理能力。

Snyder 和 Cappelleri(2001)通过计算所有测试二数分的平均数来评估受试者的表现和策略，该分数比其他传统的 PASAT 分具有更高的临床灵敏度。

Diehr 等(2003)则使用更简单的方法，要求如果受试者在测试时使用组块策略，测试者需停止测试并重复指导语。

4. 计算机分数　计算机版具有很多的评分，包括显示遗漏回答的列表、错误数、抑制错误数(受试者将呈现数字与之前所回答的数字和相加，而不是与上一数字相加)、连续三个正确回答个数、最长连续正确回答长度。但很少人知道这些分数的心理测量意义和研究相关性。电脑版也有受试者反应时间，但因为各种不同技术的限制，如不能区分实际回答和其他声音(如清喉咙)、因软件或硬件限制不能精确地反映时间等，受试者反应时间未得到应用。

调整版 PSAT 使用时间阈限(正确回答所需最短 ISI)和正确回答数评分，这可能是最能反映信息处理速度的变量。

四、人口统计学影响

1. 年龄　在大多数样本中，年龄与 PASAT 表现相关。在一些研究中，年龄的影响主要存在于 50 岁以上老人。综合所有正常人的表现发现，所有年龄段的受试者都存在 ISI 越短测试表现越差的现象。年龄和较差的数学能力影响儿童的信息处理能力。在指定时间中，年龄大的儿童能给出更多的正确回答。随着年龄增大，信息处理速度指数性地加快，尤其是 8～10 岁儿童。

2. 教育程度/IQ　教育程度与 PASAT 表现相关，教育程度越高，测试表现越好。相似的是，IQ 也与 PASAT 表现相关。然而，教育程度的影响并不见于所有研究，如 Brittain 等发现教育程度对 Levin 版 PASAT 表现几乎无影响。最近，大规模的正常人 Levin 版 PASAT 研究报道了 IQ(非教育程度)与 Levin 版 PASAT 的表现具有相关性。

3. 性别　一些研究中，男性表现略优于女性，但这影响并无临床意义。在一项大规模的研究中，性别对 PASAT 分数的影响小于 1%。无报道显示性别对电脑版 PASAT 有影响。在儿童中未发现性别影响。

4. 种族　很少有研究报道种族/文化背景对 PASAT 表现的影响。大规模的正常人 Levin 版 PASAT 研究中，Wiens 等(1997)发现种族对 PASAT 的

影响实际上是因为IQ、年龄和教育程度，而不是种族本身。Diehr等(2003)报道了年龄、教育程度和种族显著影响简短版PASAT的得分。

五、信度

1. 内部一致性　四组PASAT测试都具有很高的克隆巴赫系数($r=0.90$)。在不同年龄段的儿童中，儿童版PASAT折半信度都接近0.90，提示其具有较高的内部一致性。在青少年中，儿童版PASAT折半信度也很高($r=0.96$)。各组的测试得分也有很高的相关性(0.76～0.95)。

2. 重测信度　PASAT测试短间隔(7～10天)的重测信度很高($r>0.90$)。部分研究报道了PASAT测试长间隔(如4周)的重测信度有所降低($r=0.73$或更高)。PASAT测试长时期(超过3个月)的稳定性也很高($r=0.83～0.96$)。儿童版PASAT4周间隔的重测信度为0.78～0.83。

3. 练习效应　PASAT具有明显的练习效应。在给正常受试者PASAT测试后1周再进行测试，第二次测试得分较第一次提高18%(6分)。相似的练习效应也见于脑外伤(TBI)和HIV感染患者。在儿童中，重测得分提高约20%(平均3.8～7.3分)。Baird等发现电脑调整版PSAT的练习效应也很显著，无论第二次测试是20分钟后、1周后或3个月后。效应持续6个月，并不受刺激的感觉模式(听觉 vs. 视觉)和数字表的类型(简单和 vs. 复杂和)影响。PASAT存在高练习效应的其中一个原因是，在了解测试方式后，受试者重测时减少了其客观的紧张情绪，故受试者在基线时表现较差而重测时分数较高。毫无疑问，PASAT会引起某些受试者的焦虑情绪。

六、效度

1. 各版本间相关性　听觉模式下，电脑版和录音带版具有可比性，但仍需更多的研究证实。简短版和长版在正常人和HIV患者中均具有高相关性。视觉版和听觉版也具有高相关性。听觉版PASAT(Gronwall)和视觉版PVSAT具有高相关性($r=0.63～0.73$)。但两个版本间的基本差别可能会在不同临床患者之间产生不同的结果。同其他注意测验一起评估时PASAT和PVSAT具有相似因子结构。Fos等在正常人和脑外伤患者中未发现PASAT和PVSAT具有顺序效应；但在某些患者中存在表现差异，例如，多发性硬化患者在两个模式的测试中都表现差，而慢性疲劳综合征患者则表现好。

视觉版的干扰效应较听觉版少。一些研究表明使用相似的输入/输出模

式（听觉输入-语言输出）会导致刺激反应竞争，特别是随着 ISI 的缩短，竞争会更显著。当考虑到测试需要信息处理速度时，这个干扰就会降低 PASAT 对某些患者（如头部外伤）的敏感度，这也能解释为什么视觉版通常比听觉版简单。视觉版没有干扰效应是因为使用了不同刺激反应模式（视觉输入-语言输出）。通过各版本的比较显示，模式相比计算力更影响 PASAT 的表现；相反，因为没有干扰效应，视觉版减少了工作记忆的要求，更适用于评价信息处理速度。

一些版本的长 ISI 会出现天花板效应，这使得简短版不适合能力高的人。Fos 等发现在正常人和大学学历的头部外伤患者中，视觉版 PVSAT 的前两组测试具体天花板效应，而随着后两组测试难度的增加，未发现地板效应。

2. 与其他注意测验的相关性　PASAT 与其他需要注意分配的测验相似，也是主要评估信息处理能力，同时还包括至少 3 个组分：工作记忆、信息处理能力和信息处理速度。PASAT 与其他注意测验，如数字广度、d2 测验、连线测验（特别是测验 B）、视觉搜索和注意测验（VSAT）、Stroop 色词测验和选择反应时间任务呈现中等相关性。Sherman 等在头部外伤患者中也发现 PASAT 与韦氏记忆量表注意/不分心因子中等相关。同数字广度和韦氏记忆量表的计算部分一样，该测试在正常人中反映了注意-集中因子，但计算能力在测试中所起的作用更强。

大多数因子分析显示该测试也反映了其他注意/信息处理因子。例如，Larrabee 和 Curtiss 在多种神经心理门诊患者中发现同数字广度、数字串联和韦氏记忆之心理控制分测验一样，PASAT 反映了注意、即刻回忆和信息处理因子。Bate 等在严重脑外伤和正常人的混合样本中发现同日常注意成套测验（TEA）彩票任务（注意保持）和视觉电梯任务（注意转换）、数字广度倒背任务一样，PASAT 反映了注意-保持因子。在严重头部外伤患者中，PASAT 和符号数字转换测验反映的因子相同。在糖尿病患者中，PASAT 高度反映韦氏智力测验的注意集中因子。在精神分裂症谱系障碍患者中，PASAT 与第三版韦氏测验处理速度指数和连续操作测验的注意-保持因子具有高度的相关性。另外，PASAT 还反映连线测验 B 的执行专注因子。

该测试还评估与其他测验不同的注意功能。例如，Fos 等发现 PASAT 和 PVSAT 反映了一个与数字广度、连线测验和 Stroop 色词测验不同的因子。需要注意的是，因为 PASAT 评分的多样性，该因子受到测试方法的影响。

有研究表明不同的 PASAT 速度标志不同的处理阶段。Deary 等发现呈

现速度慢与 Rey 听觉词语学习测验中等相关，而呈现速度快与记忆功能基本不相关。在其他研究中，PASAT 的得分与记忆功能的相关度不大，并与 ISI 相关性不明确。

3. 与 IQ 相关性 研究表明 PASAT 测试与 IQ 至少呈中度相关。Crawford 等对 152 名正常受试者进行了 PASAT 和韦氏智力测验，首因分析表明 PASAT 高度反映智商，并超过韦氏智力测验的许多亚测验。许多研究也支持了 PASAT 与 IQ 的明显相关性。

4. 与计算能力相关性 许多研究发现不论是正常人还是临床患者，PASAT 与计算能力高度相关。在头部外伤患者中，数学相关测验（韦氏智力计算部分、美国标准化测验计算部分）是 PASAT 表现最显著的预测因子。该测试表现也与调整版 PSAT 计算分高度相关，可以解释 26％的变化，特别是在用听觉通道时。其他研究表明，PASAT 表现与简单加法的反应时间和最近的学校数学成绩及数学能力的自我评级有关。

计算能力对 PASAT 的贡献与反应时间有关。做加法时反应时间短的受试者拥有更长的时间来记住、回忆或预演上一个数字，而反应时间长的受试者则只有一点点时间来反应上一个数字。长反应时间的受试者在保持前一个数字并与新呈现的数字相加情况下的认知资源的干扰和竞争，与随着测试难度增加，ISI 逐渐减短的情况相似。

尽管 PASAT 所需的计算能力很基本，测试表现仍与加法任务的难度有关。Royan 和 Tombaugh 等发现计算简单的项目（和为 2～10）的正确率高于计算难（和为 11～18）的项目。但是，有研究认为呈现 2～9 范围的数字改为 2～19 范围的数字并不增加 PASAT 的敏感性。使用严格的数字表（所有和都为 2～10）可以减少（非消除）计算能力对测试表现的影响。

Ward 报道了年轻受试者完成测试比老年受试者简单，虽然年轻受试者其他注意测试的表现较好，但在 PASAT 的表现比美国成人阅读测试相匹配的老年受试者差。Ward 认为计算能力的跨代效应影响 PASAT 的表现。Ward 还注意到不是所有的个体都必须通过同一种方法（直接从记忆中检索）完成简单加法的，而是有反应时间不同的多种策略。这又需要提出计算能力在 PASAT 表现中所占角色的问题了。Crawford 等注意到如果存在计算能力的跨代效应，这能减弱（非消除）年龄对 PASAT 表现的影响。

CHIPASAT 表现与韦氏智力量表中的计算能力中等相关。因为儿童的计算能力不是那么好，在解释 CHIPASAT 的表现时需要注意。许多 8～9.5

岁儿童在测试时数手指，这提示该测试最好应用于具有合适计算能力的儿童。

七、临床研究

1. 脑外伤研究　PASAT 对轻度脑震荡的注意力评估很敏感。相比其他注意测验(包括韦氏记忆量表注意/集中指数、连线测验)，PASAT 对头部外伤(traumatic brain injury, TBI)患者的信息处理能力更敏感。该测试对严重头部外伤高度敏感，并与外伤发生时间相关。同时，大量研究未发现 PASAT 与头部外伤的严重度指标，如脑外伤后遗忘(PTA)和意识障碍(LOC)有关。在轻度脑震荡患者中，PASAT 与严重度无关；在轻度脑外伤中，PASAT 的敏感度比连续操作测验低，比连线测验和数字广度高。在外伤前能力高的患者中(如大学学历的头部外伤患者)，PASAT 对头部外伤的敏感度受限。

然而，相比其他注意测验(包括连线测验、Stroop 色词测验和数字广度)，PASAT 对持续性脑震荡后综合征(PCS)更有效，这可能因为其还需要工作记忆能力和信息处理速度。Schmidt 等报道其敏感度为 56%。Bradshaw 等报道了 45%严重脑外伤患者的 PASAT 得分在或低于 10 个百分点。总的来说，测试特异度比敏感度高，即测试对排除 PCS 和 TBI 比确诊更实用。PASAT 在一些研究中的敏感度不理想可能与各研究中参数不同有关。Bate 等发现受试者 ISI 长的组(ISI 2.4 s 和 2.0 s)不能区分严重脑外伤患者和对照组，严重脑外伤患者在 ISI 短(ISI 1.6 s 和 1.2 s)的组 PASAT 得分低。PASAT 对脑外伤的敏感度与外伤类型相关，至少部分相关。Roman 发现 PASAT 对检测伤害是由于第二次加速/减速力量和伴有皮质下损害的脑外伤最有效，如机动车事故后脑外伤。然而 Sherman 等发现 PASAT 得分与机动车事故患者的 PTA 或 LOC 不相关。

回顾性 TBI/PCS 研究可能存在选择偏移。这些研究结果取自能完成测验的患者，这说明样本可能更代表了能力高的患者，使得研究的拒绝/停止率不同。

2. 多发性硬化(MS)研究　有关 MS 患者的 PASAT 研究很多。因为大脑信息处理是以皮质下系统和脑白质连接为基础的，而这些区域在 MS 患者中优先受影响。MS 患者是信息处理速度受损而不是工作记忆受损，故 PASAT 比评估工作记忆的测试(如第三版韦氏测验的字母-数字排序)更敏感。

因为 PASAT 对 MS 相关认知损害的高敏感性，故 PASAT 被认为是评估

MS患者的核心测验。多发性硬化功能综合评分所采用的PASAT是间隔3.0 s、60项的测试，它被用于多项多中心的MS结局的研究中。同时它也是多发性硬化成套神经心理筛查量表（Rao，1991）的组成部分。

PASAT不同得分的MS患者对其的敏感度不一样。一些研究未发现MS和对照组有区别。而在使用正确数而不是二数分来评价测试表现时这些阴性结果将消失。Fisk和Archibald发现随着速度的加快，MS患者和对照组总正确数的差别逐渐减少，而反映回答策略的二数分/组块分的差别逐渐显著。换句话说，他们发现MS患者在后面的测试中跟不上测试要求时会人为地降低测试的难度。当测试难度太高时，正常人和MS患者都会有这样的表现。

PASAT对评价MS治疗效果敏感，并与损害体积高度相关，特别是在使用四组测试的二数分平均值作为指标时。PASAT对评价MS亚型也很敏感。

对于其他临床患者，PASAT对评价中度低血糖敏感，增加了反应时间和遗漏数。有研究报道慢性疲劳综合征患者的PASAT表现差，重症肌无力患者的表现并无报道。PASAT对非痴呆的帕金森病患者的认知加工速度减退非常敏感。PASAT对评价精神分裂症谱系障碍（包括首发精神障碍和分裂型人格障碍）高度敏感。PASAT还被用于评价慢性疼痛患者的认知功能和足球队员头顶球对其的认知影响。系统性红斑狼疮患者即使没有神经损害，其组块分也较预期差。

儿童CHIPASAT的实用性仍需要研究证实。Mendley和Zelko报道了肾脏移植后儿童CHIPASAT的表现有提高。

3. *神经解剖学相关性研究* 大脑定位研究表明PASAT激活了许多区域，与其多因素任务需求和其所评估的特异性认知损害一致。功能磁共振（fMRI）研究显示PASAT可激活大面积脑网络，包括注意和工作记忆网络（额顶区）；发现MS患者与对照组在评估PASAT时激活不同的大脑系统。PVSAT激活MS患者的双侧额叶（Brodmann 6、8、9、39区），而激活对照组右扣带回（Brodmann 32区）。MS患者为保持正常表现和（或）补偿功能而调动其他大脑系统被认为是导致激活区域不同的原因。

在解释PASAT fMRI研究结果时需注意输入（视觉和听觉）和输入模式（使用不需要公开回答的默想和标准的口头表达）的区别。如果使用单刺激-反应通道（听觉-语言）比双刺激-反应通道（视觉-语言）所产生的干扰更少，fMRI则不需要受试者实际回答来避免运动干扰。

4. *焦虑、任务难度、情绪和疲劳的影响* PASAT拥有一个令人厌恶的测

试的名声。它被实验性地用于引起压力和增加疲劳。Diehr 等对 HIV 患者长期随访研究发现患者宁愿做腰椎穿刺也不愿做 PASAT。研究中，超过一半的受试者在完整长度的 PASAT 测试中感到不适，而 30%的受试者在简短版 PASAT 中感到不适。在一项涉及 MS 患者的研究中，PASAT 的拒绝/停止率最高，17%的患者拒绝尝试测试，6%的患者在听指导语时放弃。另一项研究发现，PASAT 使原本情绪平稳或高兴正常的受试者情绪变差。情绪悲伤的受试者在 ISI 1.2 s 的那组测试里的表现比情绪平稳的受试者差，提示了情绪对测试表现的负面影响。事实证据表明 CHIPASAT 会使儿童紧张。笔者于 2008 年在上海城市社区做正常老年人认知功能调查，也发现 PASAT 的完成率只有 2/3 左右。

毫无疑问，PASAT 令人感到吃力和沮丧，对焦虑患者不适合。Mathias 等回顾了以前的研究发现即使是非焦虑受试者，指导任务也可能有困难。传统 PASAT 修改版(如简短版或调整版 PASAT)减短了测试长度，对减少受试者不适有帮助。还有一个版本测试(PASOT)使用单词替代数字来避免计算能力的影响，虽然受试者对其不那么厌恶，但其难度较标准版本低，可能对某些人群不敏感。

虽然临床印象是疲劳会影响测试表现，但 Johnson 等研究发现，不论是在正常人中还是在高疲劳水平的患者如慢性疲劳综合征、MS 和抑郁患者中，在同一时期测试 4 次也未发现疲劳对结果的影响，抑郁严重度与 PASAT 表现是否明显相关还有争议。Demaree 等认为在 MS 患者中，抑郁增加了 PASAT 所评估的信息处理速度损害。

5. 生态效度　PASAT 与患者对症状的主观感受、人格改变和重返工作的准备状态有关。PASAT 还是头部外伤后工作状态的强烈预测因子。在 MS 患者中，PASAT 与患者的生活质量，特别是生活质量的精神和情绪方面和模拟驾驶时事故发生率有关。

PASAT 对鉴别装病敏感。假装脑外伤的受试者的表现比未诈病者差得多，至少在间隔时间 2.0 s 的那组测试时明显。研究表明在使用 PASAT 评估有夸张风险的人群(脑外伤、脑震荡后综合征患者)时，需要采用技巧(症状有效性测验)来鉴别假装认知损害的受试者。

八、评价

作为评估受试者的工作记忆、注意分配和信息加工速度的工具，PASAT

已经被使用了很多年，它对多种临床情况敏感。它还适用于能力高的受试者，不至于会像其他注意测试一样出现天花板效应，因为PASAT需要计算能力，与IQ及教育程度相关，适用于平均水平及以上的受试者，包括能力高的儿童。最近的几个大型正常资料也增加了测试的适用性。运用时可以根据人口统计学的不同使用不同版本的PASAT，如PASAT-200、PASAT-100和PASAT-50。

测试者需要知道PASAT的一些独特的限制，在哪些情况和对哪些患者不适用。PASAT使用前需要筛查受试者的计算能力，以避免干扰。简短版PASAT(PASAT-50和PASAT-100)也很实用，但其只有一组测试，可能有明显的天花板效应，因此两组的PASAT-100更合适。

PASAT具有不同版本。版本间最显著的区别为模式不同，另外每组的测试项目数、ISIs和项目数字表也有不同。数字表不同比例的复杂加法和简单加法也会影响测试表现。所以仔细选择PASAT版本是很有必要的。

测试时还需要注意的是测试模式。多项研究表明使用相同感觉通道的输入/输出模式(听觉输入-语言输出，标准录音带版)会导致刺激-反应竞争，特别是ISIs缩短时更显著。当受试者关注听明白数字时可能来不及表达数字，关注表达数字时可能来不及听清楚。当考虑到测试需要信息处理速度时，这种干扰就会降低PASAT对头部外伤的敏感度，这也能解释为什么视觉版通常比听觉版简单。视觉版没有这种干扰效应是因为使用不同刺激反应模式(视觉输入-语言输出)。视觉版更适用于能力低的受试者，这些受试者在标准版测试时可能会出现地板效应。然而，各视觉版之间不是等价的，例如，调整版PSAT使用一个视觉遮掩来防止残留影像，而该方法在其他视觉版中未被使用。视觉版减少了工作记忆的要求，可以纯粹评估信息处理速度。

测试者在给予受试者录音带版测试时，最好使用匹配其英语口音的录音带(如给美国受试者美国口音的录音带)。然而，英语口音差别似乎不影响测试表现。有研究显示加拿大受试者无论给予美国口音还是新西兰口音的录音带，PASAT的得分相似。

分数来源一样很重要。当受试者表现很差时，错误主要为遗漏错误而不是错误回答。越来越多的证据表明传统得分(如总正确数或每个正确回答所需时间)不适合反应工作记忆和信息处理速度。正如前面所说，受试者随着测试难度的增加，会倾向使用组块策略，这时使用总正确数就会过高地评价受试者的工作记忆和信息处理速度，就需要使用二数分百分比或组块分百分比。

当前者低而后者高时，说明受试者跟不上测试的速度。PASAT－50、PASAT－100和PASAT－200通过测试者在测试开始使用组块策略时通过停止测试、重复指导语来避免这个问题，通过使用调整版PASAT可以完全消除这个问题。

另一个需要注意的评分是信息处理速度。测试者倾向通过调整版PASAT来评估受试者信息处理速度。调整版PASAT可以减少工作记忆需求而更好地评价信息处理速度。但需要注意的是，调整版PASAT根据受试者的能力调整ISI，也就消除了ISI的敏感性。

反应时间也是一个合适的“速度”指标，几乎无天花板效应，而反应时间需要通过电脑测试，其数据的可靠性受到各种软件/硬件的限制。Wingenfeld等也提到，电脑通过声音激活技术编码反应时间数据，即电脑是根据受试者第一个发声（清喉咙、叹气等）而不是第一个实际回答来收集数据的。

使用录音带版本时无疑需要消除时间准确性问题，录音带重复拷贝会降低声音质量，改变ISIs。录音带使用一段时间后，测试者特别需要检查ISIs。录音带重复使用后质量会下降，录音带重复拷贝会丧失时间精确性，所以频繁使用PASAT的测试者需要准备更换的录音带。R. McInerney建议如果录音带版本测试每天使用，录音带需每年更换2次。

数字加法：不同文化，比如中文的数字发音都是单音节，中国儿童通常要求熟背加法口诀表，强化心算、珠算，进行简单数字相加像自动回答一样熟悉，估计会影响信息处理速度，当代儿童是否习惯使用计算器也可能影响心算速度，所以，PASAT需要建立当地的、当代的常模。

最后，该测试需要快速的语言回答，这阻止了该测试应用于构音障碍或其他语言障碍的患者。该测试不仅不适用于严重焦虑患者，还会引起非焦虑患者的负面反应。该测试也不适用于正处于负面情绪的患者，ISI最快的测试对情绪的影响可以独立于注意损害而影响测试表现。一些患者可能拒绝或中断测试，测试后对患者进行解说可能可以减少该测试对情绪的负面影响。

（陈科良　郭起浩）

第七章

视觉空间能力的评估

当我们观察一个物品的时候，我们不仅想知道它是什么，还要知道它在哪里，即它的空间位置，这两方面的知识是有不同的相互独立的脑结构基础的。常用的视觉空间能力评估测验包括气球划销测验（balloons test）、钟划销测验（bells test）、Benton 面孔再认测验（facial recognition test，FRT）、线方向判断测验（judgment of line orientation，JLO）、Rey-Osterrieth 复杂图形测验（Rey-Osterrieth complex figure test，CFT）、视觉物体与空间感知测验（visual objective and space perception battery，VOSP）视觉运动整合测验、Hooper 视觉组织测验（Hooper visual organization test）、物品拼凑测验、图形排列测验、画钟测验（clock drawing test，CDT）、积木测验、图形模仿测验、顶叶测验。本节介绍 CDT、CFT、JLO 和 VOSP。

第一节　画钟测验（CDT）

画钟测验（clock drawing test，CDT）是不需要特殊材料、可以在床边进行的简短测验，所以是各种成套测验（如 CAMCOG，七分钟筛查量表，MoCA 等）的组成部分。CDT 通常有自发画钟和模仿画钟两种方式，此外，钟面时间阅读、钟面指错、图-时匹配、钟面知识测验也见诸文献报道。目前最常用的是自发画钟，通常是要求受试者在预设圆圈的白纸上画“1 点 50 分”或“3 点 40 分”的钟面。

一、评分

方法非常多，但至今没有统一的评分标准。以下介绍较有代表性的一些

评分方法。

1. 总分 3 分　评分法：画出圆形，计 1 分；钟面数字正确，计 1 分；标出正确时间，计 1 分。

2. 总分 4 分　Pfizer Inc. 和 Eisai Inc 评分法：共 4 分。画圈、正确数字、正确的数字位置及正确的指针位置各得 1 分。Mini-cog 由 3 词延迟回忆＋CDT组成，其中 CDT 的评分方法就是该 4 分法。

3. 总分 5 分　Shulman 编制的定性评分：完整 5 分，小的视觉空间错误 4 分，指针不准确 3 分，视觉空间组织中度混乱 2 分，视觉空间组织重度混乱 1 分，不能画或根本不像钟 0 分。≤3分有损害。

4. 总分 7 分　Watson 的评分方法：① 用两条线把钟面平分为 4 份，其中一条线穿过表心和数字 12，另外一条线垂直平分第一条线。② 从数字 12 的位置开始按顺时针方向计算每个象限中的数字，每个数字只能计数一次。如果某个数字恰好在两条平分线上，则这个数字归于线的顺时针方向的那个象限。正确的是每个象限有 3 个数字。③ 第一、二、三象限中的数字的数目有任何错误，计 1 分。如果第四象限中数字的数目有错，则计 4 分。④ 正常范围 0～3分；4～7 分为异常。其检测痴呆的特异度和敏感度分别为 87％和 76％。

5. 总分 10 分　Wolf-Klein 评分法：按照错误程度评分1～10分。6 分为划界分，≤6 分可考虑为认知障碍。与 Sunderland 评分法不同，Wolf-Klein 法白纸上已画好圆圈，而且也不要求时间设置。其检测痴呆的特异度和敏感度分别为 93％和 87％。

6. 总分 10 分　Sunderland 评分法。10 分：正常画钟，数字和指针在正常位置；9 分：指针位置轻微错误，如指向数字不精确，但不是偏向旁边数字，或者钟面遗漏一个数字；8 分：时针和分针更明显的错误，如偏向达到 1 个数字，或者数字位置出现间断；7 分：时针和分针显著的错误，如偏向超过 1 个数字，或者数字位置非常不恰当，如都画在一侧；6 分：指针运用不恰当，如越过数字或圆圈，或者数字云集在钟面最后或数字重复；5 分：指针不是清楚地指向数字，或者数字安排不恰当，如圆点表示数字；4 分：没有指针或画在圆圈外面，或者数字缺失或画在圆圈外面或次序错乱；3 分：没有指针，或者数字和圆圈没有关联；2 分：有理解指导语的证据，但钟模糊不清；1 分：不相干，不理解，不动笔。总分≤5 分可认为存在认知损害。检测痴呆的灵敏度为 78％，特异度为 96％。

7. 总分 16 分　Royall 的评分方法：空间位置（12～6 轴线两侧的数字对称）得 2 分，以下条目各得 1 分：总体图形像一只钟；画出圆圈；直径大于 1 寸；

所有数字在圆圈内；先画出“12，3，6，9”；假如空间位置有错误，有改正或涂擦现象；只有阿拉伯数字；呈现的阿拉伯数字只有 1～12；数字次序正确、没有遗漏或插入；只有 2 根指针；时针和分针都有箭头；时针指向正确；分针比时针长；没有以下情况：写上 1 点 50 分或文字或图片或其他东西。自发画钟(CLOX1)<10 分、模仿画钟(CLOX2)<12 分有损害。

8. **总分 20 分** Mendez 评分法，其中 3 分反映钟的一般特征；12 分反映数字及位置；还有 5 分反映指针的位置。正常≥18 分。

9. **总分 30 分** 笔者编制的一个评分系统(30-score system)：① 所有数字在钟面圆圈内得 3 分(数字逾过钟面或遗漏的扣除 0.5 分/个，直至 0 分)；② “12，3，6，9”4 个标志性数字分布对称得 2 分(3 和 9 在水平线上 1 分、12 和 6 垂直线上 1 分)；③ 其他数字的位置得 3 分，每个数字位置错误或偏移扣除 0.5 分，直至 0 分；④ 中央点位置正确得 1 分；⑤ 钟面完整得 1 分；⑥ 先锚定“12，3，6，9”4 个关键点得 4 分(每个点 1 分)；⑦ 12 个数字书写得 4 分(每个数字错误或遗漏扣除 0.5 分，直至 0 分)；⑧ 顺时针排列得 1 分；⑨ 1～12 数字次序正确得 1 分；⑩ 画时针和分针得 2 分；⑪ 时针指向正确得 2 分；⑫ 分针指向正确得 2 分；⑬ 分针比时针长得 2 分；⑭ 时针和分针都有箭头得 2 分。

其中反映画钟过程的是第 6 项，称为“CDT30 分法- A(Anchoring)”，共 4 分；反映钟面表现的是其余 13 项，称为“CDT30 分法- C(Clockface)”，共 26 分，总分 30 分。30 分法- A≤2 分，识别 MCI 的敏感性为 70.6%，特异性为 73.9%；30 分法- C≤17 分，识别 MCI 的敏感性为 38.2%，识别轻度 AD 的敏感性为 75.3%，特异性为 75.9%。

10. **错误评估** Lessig 等(2008)发现时间不正确、没有指针、数字缺失、数字重复或替换(如用符号代替数字)或拒绝画钟等 6 个错误的组合得分识别痴呆的敏感性是 88%，特异性 71%，优于 Mendez 评分系统、Shulman 评分系统和 Wolf-Klein 评分系统。

画钟测验常用的有以下 8 种形式：① 无圆圈自发画钟(常用时间点是 1:45，涉及左右两侧，并需要简单除法)；② 有圆圈自发画钟(如 11:10)；③ 有圆圈、有锚定数字的自发画钟(如8:20)；④ 模仿画钟；⑤ 系列钟面时间阅读；⑥ 钟面指错；⑦ 图-时匹配：如将“十点差一刻、10:45 等”与图面匹配；⑧ 钟面知识。

二、评价

一般认为要求受试者模仿已画好的钟反映的是非语言的空间结构能力，

这反映了右侧或双侧颞顶叶的功能；要求受试者在空白的纸上画钟，需要整合空间组织、数字次序和时间概念等多个任务，反映的是执行能力，这是额叶的功能。应用 Sunderland 10 分法编制常模制定的划界分为小于 6 分。CDT 判别 AD 的敏感性和特异性有比较大的差异，Cahn 的大样本调查结果分别是 82%和 98%，综合众多文献，CDT 的敏感度和特异度平均约 85%。CDT 对 MCI 有一定的识别作用，即它可用于预测 AD 的发生。血管性痴呆患者在书写运动、数字安排、指针位置和执行控制错误（如标数字时转动纸张、逆时针标数和持续动作）方面的总错误数多于严重度匹配的 AD 组。有的评分标准中要求在钟面上优先标出“12，3，6，9”这 4 个数字，中国小学以上教育程度的健康老人中运用这一策略使两侧数字对称分布的比例大约是 80%。

第二节　复杂图形测验(CFT)

复杂图形测验（complex figure test，CFT）最早在 1941 年由 Rey 设计而成，当时作为一种评定脑损伤患者视空间结构能力及视觉记忆的工具。1944 年 Osterrieth 用自己的评分系统将 Rey 的施测程序标准化并首次提供了 230 例儿童和 60 例成人的常模，此后便成为现今所广泛应用的神经心理测验——Rey-Osterrieth 复杂图形（Rey-Osterrieth complex figure，ROCF）测验（图7-2-1），在临床工作及研究中用于评定视觉空间结构能力和非语言记忆。

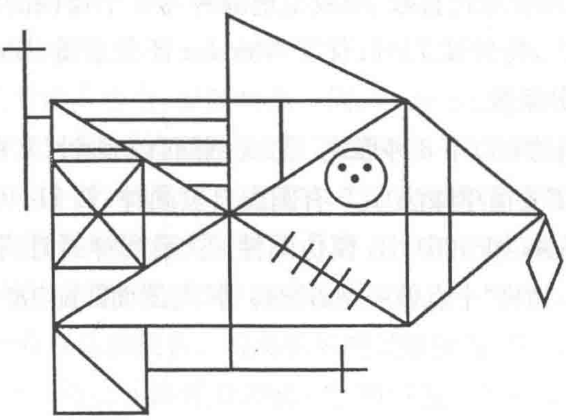


图 7-2-1　Rey 复杂图形

尽管诸如 Taylor 图(图 7-2-2)、MCG 图(图 7-2-3)等被设计成可与 Rey-O 进行比较并用于复测，Rey-O 仍是使用最广泛的复杂图形。由于 Taylor 图的得分高于 Rey-O 图，Hubley 和 Tremblay(2002)减少了 Taylor 图中一些有特点的部分，附加一些线条以增加视觉测验的复杂性，并对一些有特点的细节位置做调整。

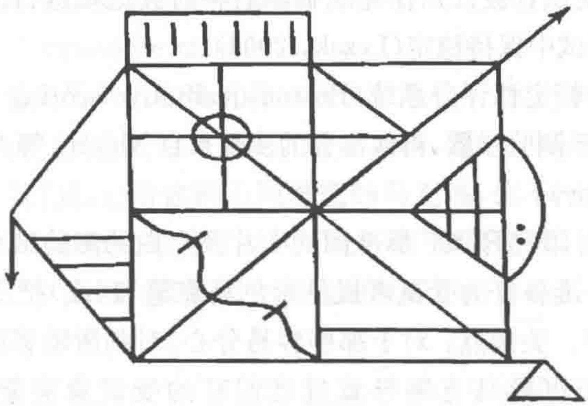


图 7-2-2　Taylor 复杂图形

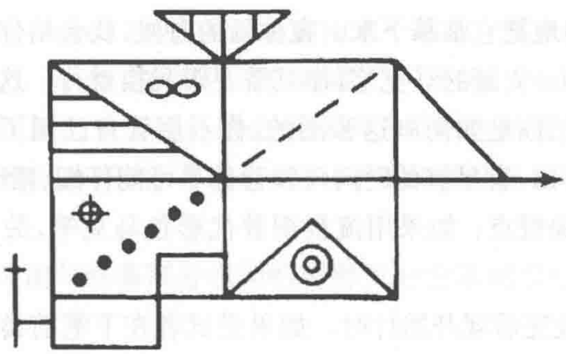


图 7-2-3　MCG(Medical College of Georgia)复杂图形

一、测验材料

(1) 黑色或彩色钢笔，铅笔或彩色马克笔(felt-tip markers)。

(2) 计时器。

(3) ROCF 标准图(8 英寸×11 英寸)及评分手册。

说明：出版的手册包括用于记录各项原始分及标准化分的分数汇总页，以及其他用于记录即刻回忆、延迟回忆和(或)再认的部分。另外，还可以为测

试者提供素描图用于跟踪受试者描画次序并做相关记录。

二、测验步骤

在大部分施测程序中，说临摹部分的指导语时不事先告知受试之后将要求其根据记忆重新画出该图。MCG 图被设计用于临床试验的重复测验，复测时间间隔较短，受试者被告知在完成临摹后将对其记忆进行测试以保证该任务要求在整个测试中保持稳定(Lezak, 2004)。

以下以波士顿定性评分系统(Boston qualitative scoring system, BQSS)为例，介绍 ROCF 测验步骤，再认部分的步骤来自 Meyers 等人的研究。

1. 临摹(copy)

(1) BQSS 将印有 ROCF 标准图的卡片及空白的测验纸水平地放在受试者面前的桌上。准备好为受试者提供彩色马克笔和(或)把受试者的画同步记录在流程图中。关键点：对于那些容易分心、对刺激敏感或那些有其他临床障碍可能由于更换马克笔导致过度偏移的受试者应避免使用彩色马克笔。

(2) 指导语："请你临摹这张图片。不需要像艺术家作画那样，只需要你尽可能仔细、精确地把它临摹下来。在你画的时候，我会给你不同颜色的马克笔，每当我递给你一支新的马克笔，你只需要跟我换就行。这些彩色马克笔只是用来帮助我记忆你是如何画这张图的，你不用管自己用了哪些颜色。临摹没有时间限制，所以，有足够的时间可以让你尽可能仔细、精确地临摹。现在，你可以开始了。"关键点：如果用流程图替代彩色马克笔，关于彩色马克笔的指导语可省略。

(3) 指导语说完后就开始计时。如果受试者在下笔前尝试旋转刺激卡片或空白测验纸，请将其恢复至水平方向。但如果受试者在已开始临摹之后旋转刺激卡片，则无须干预，将其旋转的情况做记录即可。

(4) 记录笔画顺序。这步可通过使用彩色马克笔或流程图两种方法完成。可选择其中一种或两者兼用。当用彩色马克笔时，在下列情况需更换颜色：开始画一个细节(element)，尚未完成就去画另一个细节；作画中有停顿；将图中的细节分解；犯其他相关的计划错误(Stern 等人指出按照事先决定的时间间隔，如 30 s，换笔效果不佳)。不允许在受试者画线条的中途或画好部分的局部细节时换笔而迫使其图形破碎。使用流程图时，无论是替代彩色马克笔还是同时使用彩色马克笔都要在受试者临摹时同步记录其绘图情况。彩

色马克笔的使用提供了丰富的关于笔画的视觉记录，但流程图可能在描绘笔画顺序方面比彩色马克笔更精确和完整。具体采用何种方法一部分取决于受试者的性格特征，一部分取决于测试者的偏好。

（5）当受试者完成任务后，记录临摹所花的时间，移去标准图和受试者画的图。

（6）不要告诉受试者之后还会要求其根据记忆重画。

2．即刻回忆（immediate recall）

（1）在临摹结束后马上在受试者面前水平地放一张空白的测验纸。

（2）在桌上放完测验纸后马上说下述指导语："接下来请你根据记忆重新画刚才那幅图。我仍然会给你换不同颜色的马克笔，你不用顾虑用了哪些颜色。再次声明，你有足够的时间可以尽可能仔细、精确地画。请开始。"

（3）开始计时并采用与临摹部分相同的可辨识的方法同步记录受试者的作画顺序。

（4）当受试者完成这部分任务后，按照要求记录所花的时间以结束延迟回忆部分，移去受试者画的图。

3．延迟回忆（delayed recall）

（1）等待 20～30 min 后，在受试者面前水平地放一张空白的测验纸。

（2）指导语："一小会儿前我曾请你临摹了一张图。现在请你再一次根据记忆把刚才那幅图画出来。我仍然会给你换不同颜色的马克笔，你不用顾虑用了哪些颜色。再次声明，你有足够的时间可以尽可能仔细、精确地画。请开始。"

（3）计时并采用与临摹部分相同的可辨识的方法同步记录受试者的作画顺序。

（4）当受试者完成这部分任务后，记录延迟回忆所花的时间并移去受试者画的图。

4．再认（recognition）

（1）在延迟回忆部分结束后，将再认测验纸及笔放于桌上。

（2）指导语："这张纸上的一些图案是之前请你临摹的那张大图上的局部。请圈出属于你所临摹的那张大图局部的图案。这张图上的每个图案的方向均与原始完整的图一致。共有 4 页，图案的编码从 1 到 24。请开始。"

5．不同的施测程序　4 种不同的施测程序见表 7-2-1。临摹是必须的步骤，不同版本的回忆时间间隔不一样。

表 7-2-1　4 种不同的施测程序

施测程序	A	B	C	D
临摹	√	√	√	√
即刻回忆	—	√	—	—
3 min 回忆	—	—	√	√
30 min 延迟回忆	√	√	—	√
再认	—	—	—	√
评分标准	LB Taylor	EM Taylor	EM Taylor	Meyers and Meyers

三、测验时间

（1）测验时间：40～60 min。

（2）临摹：按受试者临摹的速度而异，但一般在 5～10 min 完成。

（3）即刻回忆和延迟回忆：按受试者的绘图速度而异，但一般 5～10 min。

（4）即刻回忆与延迟回忆之间的间隔：20～30 min。

四、评分

由于 Rey 在其最初的测验描述中省略了评分标准，后人在各自的研究中采用的评分系统各不相同，既有相对较宽松的，也有相当严格的。尽管根据不同的评分系统得到的分数有所不同，主要还是包含 3 个部分：位置、准确性和结构。以下列出了一些国际上较常用的评分系统以供参考。

（一）Rey 复杂图形评分系统

1. 评分单位

（1）长方形外面左上角的十字形。

（2）大的长方形。

（3）交叉的对角线。

（4）大长方形 2 的水平中线。

（5）大长方形 2 的垂直中线。

（6）在大长方形 2 内左侧的小长方形。

（7）在小长方形 6 上的一条线段。

（8）在大长方形 2 内左上方的 4 条平行线。

（9）在大长方形外右上方的三角形。

（10）大长方形 2 内、9 下面的小垂直线。

(11) 大长方形 2 内的带三点的圆圈。

(12) 大长方形 2 内右下方对角线 3 上的 5 条平行线。

(13) 与大长方形 2 右侧相连的三角形。

(14) 与三角形 13 相连的菱形。

(15) 三角形 13 内的垂直线，与大长方形 2 的右侧垂直线相平行。

(16) 三角形 13 内的水平线，即大长方形内水平中线 4 的延续。

(17) 大长方形 2 下面的十字形，与大长方形的垂直中线 5 相连。

(18) 与大长方形 2 左下方相连的正方形。

2. **评分方法**　对 18 个单位分别进行评分，每个单位最高分为 2 分，总分为 36 分。评价每个单位的准确度及在整个图中的相对位置的细则如表 7-2-2 所示。

表 7-2-2　Rey 复杂图形测验评分标准

图　形	位　置	得分
图形正确	位置合适	2
	位置不合适	1
图形扭曲或不完整，但可辨认	位置合适	1
	位置不合适	0.5
图形缺失或不可辨认		0

3. **评分意义**　临摹低分说明视觉感知、视觉运动整合能力下降，即刻及延迟回忆低分提示视觉空间回忆能力下降。此外，通过进一步分析可得到编码模式及储存过程方面的信息。比如即刻回忆高分但延迟回忆低分说明记忆储存过程被打乱。

4. **评分系统举例**

(1) Rey-Osterrich 复杂图形的评分：

1) 测验图形：见图 7-2-4。

2) Taylor 对 Rey-Osterrich 复杂图形的评分单位：见表 7-2-3。

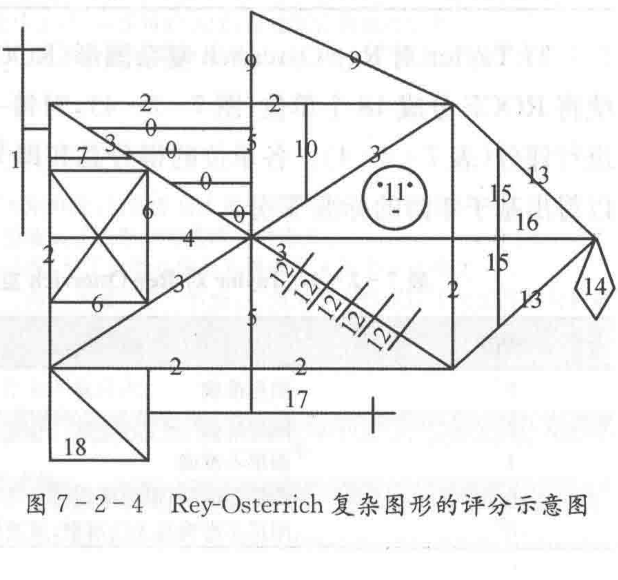


图 7-2-4　Rey-Osterrich 复杂图形的评分示意图

表 7-2-3　Taylor 对 Rey-Osterrich 复杂图形的评分单位

1. 左上角的十字，在长方形之外，与长方形的平行，其位置必须在长方形上方，连接十字与长方形的线必须接近十字的中点，介于第 7 条和长方形的顶边之间。
2. 大长方形，长方形的水平长度不必超过垂直长度的两倍，但不能像正方形。由于长方形变形的可能性非常多，对位置进行评分是不可能的，如果长方形不完整或有任何形式的变形，只给予 1 分。
3. 对角线必须连接长方形的 4 个角，相交在长方形的中心。
4. 长方形的水平中线须以一连贯的直线清晰地通过长方形左边中点与右边中点。
5. 垂直中线必须起始于长方形底边中点，以一连贯的直线，穿过长方形顶边中点。
6. 小长方形必须在大长方形之内，并在其左边。小长方形的界限是长方形顶端介于直线 2 和 3 之间(这些平行线组成第 8 部分)，小长方形的宽必须接近大长方形宽的 1/4，也就是说，小长方形的右边应该通过大正方形左边与垂直中线之间的中点。小长方形内的对角线必须经过它的 4 个角，且相交于小长方形的中点。
7. 小长方形上的直线，必须短于小长方形的宽度，必须介于小长方形的顶端与第 8 条的第 2 条线之间。
8. 在长方形左上角内的 4 条平行线，必须是平行的，间距应该大致相等。如果直线过度弯曲、少于或多于 4 条，应扣分。
9. 长方形右上方的三角形，高度小于底边。
10. 长方形内的短直线，刚好在三角形下方，线条必须清晰地伸至大长方形内右上四边形的左侧。
11. 有三点的圆，必须位于右上四边形的右下部分。圆不触及三角的任何一条边，点的位置应该是两点在上、一点在下，看起来像一张脸。
12. 5 条平行线，经过右下方对角线，必须都在右下方的四边形内，它们不能碰到四边形的任何一条边，彼此之间的间距大致相等。
13. 大长方形右侧的三角形，其高不超过长方形中间水平线的 1/2，其斜边并不是第 9 条三角形斜边的延续。
14. 钻石形，连接在第 13 条的尾端。
15. 第 13 条三角形内的垂直线，必须是平行于第 2 条大长方形的右侧垂直线，在第 13 条内靠近左侧。
16. 第 13 条内的水平线，是第 4 条向右的延续，必须起自大长方形右边的中点直到三角形的顶端。如果第 13 条三角形轻微倾斜或第 4 条不通过长方形右边的中点，只要它从长方形右边中点到三角形顶端，第 16 条依然计满分(2 分)。
17. 与长方形下侧中间区域相连的十字形，其右边必须明显长于左边，但不超过大正方形的右侧，其左侧末端应在正方形右边中点上。
18. 第 2 条的左下方，必须明显是一个正方形，与第 6 条的长方形相对，它的边长必须与第 6 条长方形的宽度相等，即等于大长方形长边的 1/4。

3) Taylor 对 Rey-Osterrich 复杂图形(ROCF)的定量评分：Rey 的评分系统将 ROCF 分成 18 个单位(图 7-2-4)，对每一个单位的准确性和位置同时进行评分(表 7-2-4)。各单位的得分总和即为原始分(0.0～36.0)，同时可以得出基于年龄的标准 T 分。

表 7-2-4　Taylor 对 Rey-Osterrich 复杂图形的定量评分

得分	准　　确	位　置
2	图形准确	位置正确
1	图形准确	位置不正确
1	图形不准确	位置正确
0.5	图形不准确，但可辨别	位置不正确
0	图形不准确且无法辨别，或遗漏	位置不正确

(2) Rey-Osterrich 复杂图形(B 型)的评分

1) 测验图形：见图 7-2-5。

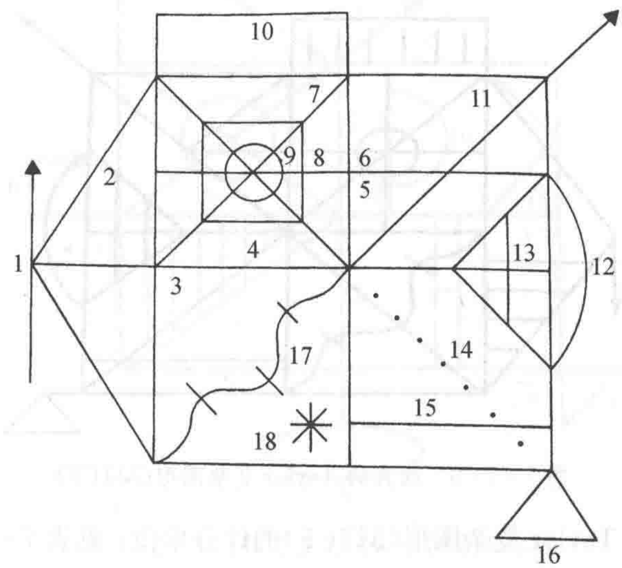


图 7-2-5　Rey-Osterrich 复杂图形(B 型)的评分示意图

2) Rey-Osterrich 复杂图形(B 型)的评分单位：见表 7-2-5。

表 7-2-5　Rey-Osterrich 复杂图形(B 型)的评分单位

1. 左侧垂直向上的箭头线，在正方形之外，与正方形平行，其位置必须在正方形左侧中间。
2. 大正方形左侧的三角，位于正方形外面，三角形底边是大正方形左侧边，顶点位于 1 中点。
3. 大正方形。
4. 正方形的水平中线须以一连贯的直线清晰地通过正方形左边中点，右边中点及 2 的顶点。
5. 垂直中线必须起始于正方形底边中点，以一连贯的直线，穿过正方形顶边中点。
6. 正方形上半部分水平中线，须以一连贯的直线清晰地通过正方形上半部分左边中点与右边中点。
7. 正方形左上象限内的对角线。
8. 大正方形左上象限内的小正方形。
9. 大正方形左上象限内的圆。
10. 大正方形左上方的矩形，在大正方形外面，且长度是大正方形的一半。
11. 大正方形右上象限外的箭头线，沿着大正方形的对角线延伸出来。
12. 大正方形右侧的半圆，在正方形外面，与正方形右侧垂直线的 1/4 及 3/4 点相交。
13. 大正方形内右侧的三角形，包括三角形内的垂直中点线，三角形 3 点分别位于大正方形右侧垂直线的 1/4 点及 3/4 点以及大正方形中间水平线右 1/4 点处。
14. 大正方形内右下象限内的点线，包括 7 点。
15. 点线之间的水平线，位于大正方形右下象限内。
16. 大正方形下的三角形，位于大正方形外，三角形底线与大正方形水平线平行，顶点位于大正方形右下角。
17. 大正方形左下象限内的曲线和正交线。
18. 大正方形左下象限内星形。

(3) 改良的 Taylor 复杂图形(MTCF)评分

1) 测验图形：见图 7-2-6。

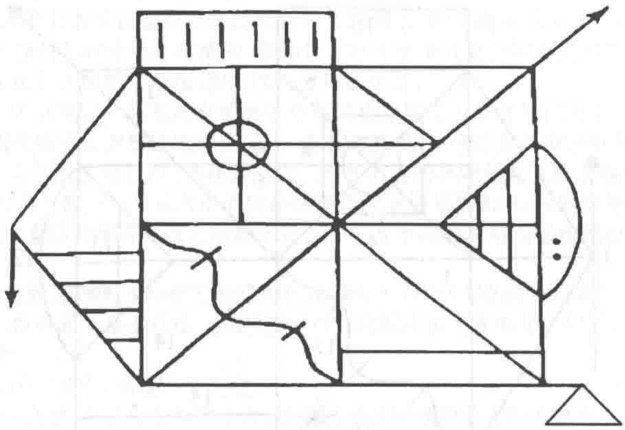


图 7-2-6　改良的 Taylor 复杂图形(MTCF)

2) 改良的 Taylor 复杂图形(MTCF)的评分单位：见表 7-2-6。

表 7-2-6　改良的 Taylor 复杂图形(MTCF)的评分单位

1. 大正方形。
2. 大正方形内的对角线。
3. 正方形的水平中线须以一连贯的直线清晰地通过正方形左边中点及右边中点。
4. 垂直中线必须起始于正方形底边中点，以一连贯的直线穿过正方形顶边中点。
5. 大正方形右上象限内的短水平线，起始于正方形右上象限左边中点及右上象限对角线的中点。
6. 大正方形右上象限内的短对角线。
7. 大正方形右角上的斜箭头线，位于大正方形外侧，是 2 的向外延伸。
8. 大正方形内右侧的三角形，包括三角形内的垂直中点线，三角形 3 点分别位于大正方形右侧垂直线的 1/4 点及 3/4 点以及大正方形中间水平线右 1/4 点处，包括三角形内的 2 个垂直线。
9. 大正方形右侧的半圆，在正方形外面，与正方形右侧垂直线的 1/4 及 3/4 点相交，包括半圆内 2 个点。
10. 大正方形下的三角形，位于大正方形外，三角形底线与大正方形水平线平行，顶点通过短水平线与大正方形右下角相连。
11. 右下象限内水平线。
12. 大正方形左下象限内波浪线，包括 2 个短线。
13. 大正方形左侧的三角，位于正方形外面，三角形底边是大正方形左侧边。
14. 13 内的 4 条短水平线，位于 13 下方。
15. 连接 13 顶点的向下垂直箭头线。
16. 大正方形左上象限内的水平中线和垂直中线。
17. 左上象限内的圆形，以大正方形左上象限中点为圆心。
18. 大正方形左上方的矩形，在大正方形外面，且长度是大正方形的一半，包括矩形里面 6 条垂直线。

(4) MCG 复杂图形Ⅰ的评分

1) 测验图形：见图 7-2-7。

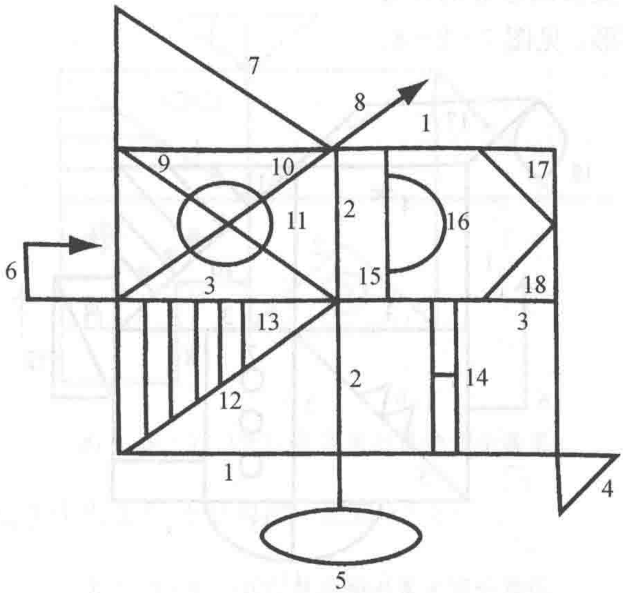


图 7-2-7 MCG 复杂图形Ⅰ评分示意图

2）MCG 复杂图形Ⅰ的评分单位：见表 7-2-7。

表 7-2-7 MCG 复杂图形Ⅰ的评分单位

1. 大长方形。
2. 大长方形内的中间垂直线。
3. 大长方形内的中间水平线。
4. 大长方形右下角外的直角三角形，其中两边分别与水平线及垂直线平行。
5. 大长方形下的椭圆形及连接线，椭圆形长轴与水平线平行，连接线连接椭圆上中点及大长方形底边中点。
6. 大长方形左侧的弯曲箭头线，起始点位于大长方形左边中点，呈 90°向上弯曲，然后再向右侧呈 90°弯曲。
7. 大长方形左上象限上的直角三角形，位于长方形外面，底边是大长方形顶线左半部，另一直角线是大长方形宽的一半。
8. 大长方形上的斜箭头线，起始于大长方形顶边的中点。
9. 大长方形左上象限内的对角线。
10. 大长方形左上象限内的另一个对角线。
11. 大长方形左上象限内的圆形，以大正方形左上象限中点为圆心。
12. 大长方形左下象限内的对角线。
13. 在长方形左下象限内的 5 条垂直线，必须是平行的，间距应该大致相等。
14. 大长方形右下象限内的 2 条垂直线及其内的水平连接线。
15. 大长方形右上象限内的垂直线。
16. 15 右侧的半圆形。
17. 大长方形右上角的对角线，起点位于右上象限右侧边中点，17 与大长方形顶边相交形成等腰直角三角形。
18. 起点位于右上象限右侧边中点，延伸到 3 的对角线，形成等腰直角三角形。

(5) MCG 复杂图形Ⅱ的评分

1) 测验图形：见图 7-2-8。

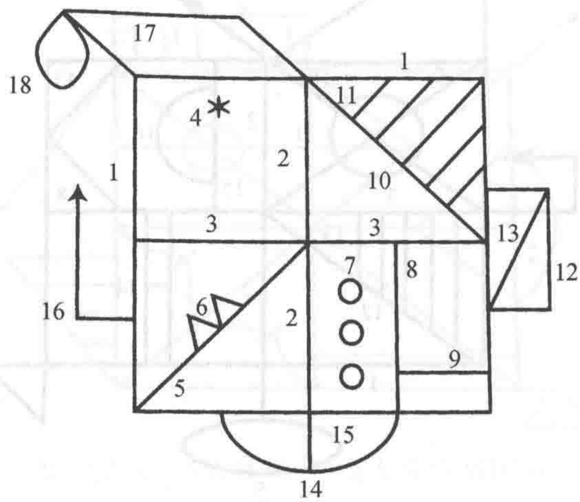


图 7-2-8　MCG 复杂图形Ⅱ评分示意图

2) MCG 复杂图形Ⅱ的评分单位：见表 7-2-8。

表 7-2-8　MCG 复杂图形Ⅱ的评分单位

1. 大正方形。
2. 大正方形内的中间垂直线。
3. 大正方形内的中间水平线。
4. 大正方形左上象限内的星号。
5. 大正方形左下象限内的对角线。
6. 5 上的 2 个紧邻的小等腰三角形。
7. 大正方形右下象限内的 3 个圆形，呈垂直排列。
8. 大正方形右下象限内的中间垂直线。
9. 连接大正方形右边到 8 的水平线。
10. 大正方形右上象限内的对角线。
11. 5 个位于 10 上方，并垂直于 10 的对角线，呈等距排列。
12. 大正方形右侧的小长方形，位于大正方形外。
13. 12 内的对角线。
14. 大正方形下的半圆形，在正方形外面，与正方形底边的 1/4 及 3/4 点相交。
15. 14 内的垂直线，起于大正方形底边的中点。
16. 大正方形左侧拐角箭头线，起于大正方形左下象限左边中点水平线，然后呈 90°向上弯曲的箭头线。
17. 大正方形上方的平行四边形。底边是大正方形顶边左半部分。
18. 连接 17 左顶部的泪滴状。

(6) MCG 复杂图形Ⅲ的评分

1) 测验图形：见图 7-2-9。

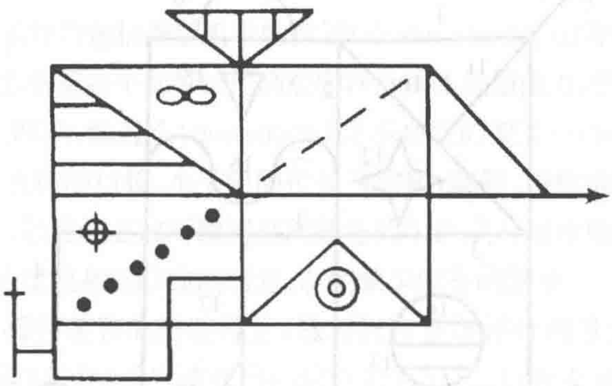


图 7-2-9 MCG 复杂图形Ⅲ评分示意图

2）MCG 复杂图形Ⅲ的评分单位：见表 7-2-9。

表 7-2-9 MCG 复杂图形Ⅲ的评分单位

1. 大长方形。
2. 大长方形内的中间垂直线。
3. 大长方形内的中间水平线。
4. 大长方形左上象限内的对角线。
5. 在长方形左上角内的 3 条平行线，必须是平行的，间距应该大致相等，3 条水平线延伸到 4 的左侧。
6. 大长方形左上象限内 4 上方的无穷大符号。
7. 大长方形左下象限内的小圆形及穿过圆心的垂直十字形。
8. 大长方形左下象限内的 6 个点对角线，6 个点大致等距。
9. 大长方形左下象限内小长方形。
10. 大长方形下面向下延伸的小长方形。
11. 大长方形左侧外面十字形通过水平短线连接于 10 左边中点。
12. 大长方形右下象限内 2 条线形成直角，以大长方形底边右半部分为底边形成直角等腰三角形。
13. 12 形成的三角形内 2 个同心圆。
14. 大长方形右上象限内虚线形成的对角线，5 条虚线大致呈等距。
15. 大长方形上的三角形，顶点位于大长方形顶边中点，三角形底边呈水平。
16. 15 内的 3 条垂直线，中间线位于 15 中间，另 2 条垂直线大致呈等距。
17. 大长方形右侧的三角形，直角边分别为大长方形右边上半部分及 3 的延长线。
18. 连接 17 的水平箭头线。

（7）MCG 复杂图形Ⅳ的评分

1）测验图形：见图 7-2-10。

2）MCG 复杂图形Ⅳ的评分单位：见表 7-2-10。

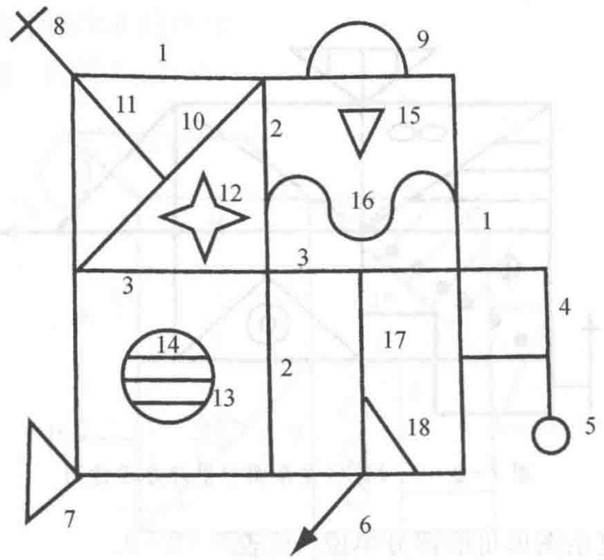


图 7-2-10　MCG 复杂图形Ⅳ的评分示意图

表 7-2-10　MCG 复杂图形Ⅳ的评分单位

1. 大正方形。
2. 大正方形中间垂直线。
3. 大正方形中间水平线。
4. 大正方形右侧的长方形，位于大正方形外面，与大正方形平行。
5. 连接 4 的圆形及其垂直连接线。
6. 大正方形底部的斜箭头线，起始于大正方形右下象限底边中点，箭头呈 45°向左下方向。
7. 大正方形左下角连接的小三角形，顶点位于大正方形的左下角，底呈垂直线。
8. 大正方形左上角外面连接的十字形，与大正方形顶边呈 135°。
9. 大正方形上方的半圆形，圆心位于大正方形顶边右半部分中点。
10. 大正方形左上象限内的对角线。
11. 10 上方与 10 垂直的延伸对角线。
12. 大正方形左上象限内的四角星，位于 10 下方。
13. 大正方形左下象限内的圆形，以左下象限中点为圆心。
14. 13 内的 3 条水平线，大致呈等距。
15. 大正方形右上象限内的顶向下的三角形。
16. 大正方形右上象限内的正弦波。
17. 大正方形右下象限内的中间垂直线。
18. 起始于大正方形右下象限底边右 1/4 点，延伸到 17 右侧的对角线。

(二) BQSS 评分系统

BQSS 评分系统提供了 17 个维度的量化分和 6 个基于这些维度的总分。这 17 个维度包括存在(presence)、准确性(accuracy)、位置(placement)、碎片整理(fragmentation)(低分说明将图片分解成零碎的部分、在细节整合方面表现很差)。存在的评价包括构型存在(configural presence)、组合存在(cluster presence)及细节

存在(detail presence)。准确性评价包括细节构型(configural element)和细节串联(clusters)。位置的评价包括图形组合的位置(cluster placement)和细节位置(detail placement)。计划,根据各个细节的作画顺序评价整体规划能力;整洁度,所画的图中有多少波浪线、间断、画过头(overshoots)及不必要的交叉(cross-outs);尺寸变形的评价包括垂直向的拉长、水平向的拉长及缩短;旋转、持续地重复、虚构作为图形的附加评价,如受到之前所做视空间任务的干扰或对图片做新的添加;不对称,对变形程度的比较和(或)在图的左、右侧缺少应有的细节。

6 项总分包括存在和准确性得分,是对损害整体评价的量化得分;即刻保持,代表临摹到即刻回忆丢失或获得的信息的百分比,以存在和准确性得分为基础;延迟保持,代表从即刻回忆到延迟回忆丢失或获得的信息的百分比;组织、评价碎片整理和计划能力,提供了更具普遍性的组织能力的评价。

(三) DSS 评分系统

DSS 评分系统可得到 3 项分数(组织、准确性和错误)及在 3 种情况下对风格(style)的评级。基于 ROCF 标准得出的组织分分值为 1～13,高分说明受试者的画具有高度的组织性及准确性。通过组织方面的得分可以确定基础组织水平(basal organization level),即儿童视觉组织能力的发育水平。

风格得分来自儿童所画的图中所表现的加权(weighting)特点,依赖于基础组织水平。有 3 类主要分风格(部分定位、中间及构型),中间类别内的 2 个子类别为我们提供了一般规律性的假设——儿童的认知风格偏向广阔的视野甚于细节,抑或相反。一般来说,部分定位(part-orientation)在年幼的儿童中更为突出。随着年龄的增长,儿童的绘图逐步显现其格局。此外,风格评级为发育趋势提供了有用的指示。

准确性得分是计算符合儿童所画的 ROCF 图中具有代表性的 64 个线段单位的数目。64 个单位按照图形类别分组,包括由大长方形和主要的子结构得出的结构性元素的准确性得分,以及由外部构型和内部细节得出的附加元素的准确性得分。在临摹部分,准确性得分表明儿童在组织、监控及材料复制方面的表现良好程度。在回忆部分,随着年龄增长,儿童对结构性元素的回忆相对优于附加元素,这说明感知基本组织框架的能力逐步发育完善,通过运用这一框架使回忆更容易。

错误总分即所犯的错误数——旋转性错误、持续性重复错误、位置错误及混为一谈的错误总数。高分可以提示发育不良的元认知策略,以及对图形中的各个元素如何相互联系存在困惑。在回忆部分,错误分高可能提示材料在

记忆中被扭曲。

（四）Savage评分系统

Savage评分系统可得到结构准确性及组织策略两项分数。结构准确性使用Denman(1987)改进的评分系统进行量化，该评分系统对图片的24个部分进行评估，使用的评分标准诸如象限位置、线的角度、线长及线条的数目。每个部分有3条标准，每条标准1分。得分范围为0～72分，高分提示良好的结构及视觉细节回忆能力，而非将各个部分组织得有多好。

组织策略包括对组织序列的量化分析和描述性分析。量化分析有5项评价（图7-2-11）——大长方形、对角线、垂直中线、水平中线及右上方顶部的三角形；大长方形作为重要的图形基本结构，评2分，余下4项各1分。分值为0～6分，高分提示良好的组织能力。在对组织序列的描述性分析中，需要注意那些从细节开始图形构建的受试者。所谓细节即5项基本构型单位以外的部分。在这样的情况下，对接下来的2个构造单元进一步检查以评估受试者是否最终能画出5个构型单位之一。这为早期组织序列提供了信息。

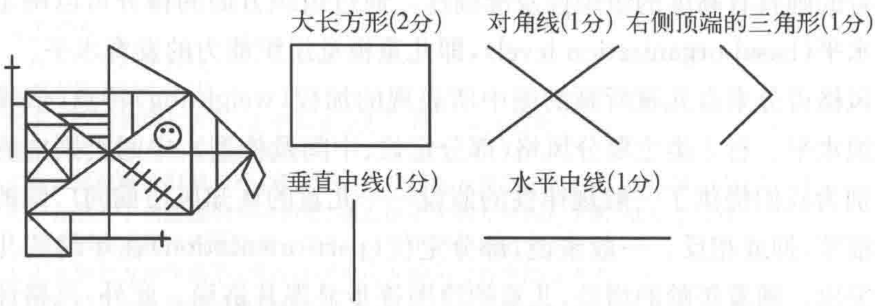


图7-2-11　Savage评分系统的5项基本构型单位

五、常模

国内汤慈美等人曾对116名正常人用Rey图及Taylor图进行测查，结果见表7-2-11。

表7-2-11　正常人Rey图和Taylor图测验结果

项　目	例数	得分		时间(s)	
		临摹	回忆	临摹	回忆
Rey图	62	35.4(0.9)	22.9(6.3)	190.5(100.3)	201.1(75.7)
Taylor图	54	35.6(0.7)	28.6(5.0)	198.7(106.9)	198.3(93.6)

笔者2000年采用Taylor(1981)的评分标准对111名正常老人的Rey图测验进行评定，该人群的模仿得分为31.2±4.0分，延迟回忆得分为14.7±5.6分，再认得分为18.9±3.4分，以文化程度进行分层得到的3个部分的均分见表7-2-12。

表7-2-12 不同教育程度正常老人Rey图临摹、回忆和再认得分

教育程度	模 仿	回 忆	再 认
文盲(n=11)	22.11(7.55)	6.33(5.87)	16.17(5.81)
小学(n=26)	29.06(5.60)	11.54(5.70)	18.72(1.81)
中学(n=50)	31.02(3.42)	15.64(5.87)	19.41(2.06)
中学以上(n=24)	33.01(2.13)	18.53(5.35)	20.20(2.03)

文盲组：未接受过教育；小学组：教育程度为1～6年；中学组：教育程度为7～12年；中学以上组：教育程度在12年以上。n为例数。

笔者2011年采用Taylor(1981)的评分标准对636名初中及以上教育程度的正常中老年人的Rey图测验进行总结，见表7-2-13。值得注意的是，与年龄、教育匹配的西方老年人相比，中国老人的CFT模仿得分高约2分，这可能与汉语文字结构和书写方式有关。英语是字母系统，汉字是图形性文字，讲究笔画整齐、间架均衡、轮廓方正，教育程度较高的中国老人，长期使用方块文字，这种训练使其在模仿画几何图形方面具有一定的优势。

表7-2-13 不同年龄正常老人Rey图临摹和回忆得分

年 龄	模 仿	回 忆
50～59(n=210)	34.41(1.63)	19.42(7.07)
60～69(n=215)	33.42(2.33)	15.95(5.20)
70～79(n=211)	33.48(3.23)	13.66(5.06)

n表示例数。

六、信度与效度

1. 内部一致性 Rey图内部一致性的评定是把每个细节作为一个子项并计算折半信度系数(split-half)及α系数。成人临摹部分的折半信度系数与α系数均大于0.60，回忆部分均大于0.80，提示所有的细节之间的联系为单因素。

2. 重测信度 Meyers等指出一些分数(如临摹、再认)的得分范围在大部

分正常受试者中，因其表现为满分或接近满分而受到限制，人为地降低了重测相关系数的大小。而且，当受试者在原来的测试结束后进行复测会使偶然学习范式(incidental learning paradigm)受到污染。基于上述考虑，Meyers 等人仅对得分偏低的受试者进行重测信度的评定，12 位正常受试者间隔 6 个月的重测信度为：即刻回忆，相关系数 $r=0.76$；延迟回忆，$r=0.89$；再认正确总数，$r=0.87$。

为了评估临床解释的时间稳定性(temporal stability)，Heaton 等对这些测量结果按年龄进行修正并得到相应的 T 分：45～54 中等，40～44 低于中等，35～39 轻度受损。对上述 3 项测试结果(即刻回忆、延迟回忆和再认正确总数)的临床解释，首次和第二次测试的百分比一致性(the percentage agreement)高达 91.7。

Berry 等对老年受试者 1 年之后进行的复测结果并不理想。他们发现在 1 年的间隔下，临摹部分的重测信度不可靠($r=0.18$)；即刻回忆($r=0.47$)和延迟回忆($r=0.59$)的系数也很低。其他研究者报道了稍微好些的结果。Mitrushina 和 Satz(1991)对老年人进行了为期 3 年，每年一次的评定，得到的信度系数临摹部分为 0.56～0.68，3 min 延迟回忆为 0.57～0.77(Mitrushina and Satz，1991)。Levine 等 2004 年亦有类似的报道。

3. *学习效应* Levine 等在 4～24 个月不等的时间间隔(平均天数=251 天，SD=129)对 478 名健康男性(平均年龄=42.2 岁，SD=8.6，平均受教育年数=16.4 年，SD=2.3，大部分为高加索人)的 Rey 图即刻回忆和延迟回忆做了评定，分别得出临摹、即刻回忆、延迟回忆的回归方程，重测的时间间隔长度和受试者的教育程度均非第二次表现的预测指标。

4. *评分员间信度* 18 项关于评分员间信度的研究得到的结果好坏不一，差的仅为 0.14，好的则高达 0.96，这意味着 Osterrieth 评分系统将受益于更为精细的量化评分标准。其他采用严格评分标准的研究得到的评分员间信度同样较高(>0.90)。

5. *效度* 相关分析及因素分析方面的研究支持复杂图形测验作为视觉构造能力和记忆(回忆及再认)的测量工具的有效性，而在特定的执行功能方面相关的数据较少。

七、临床评价

为什么语言类的记忆，如听觉词语学习测验、逻辑记忆测验，都是学习后

告诉受试者一段时间后要求再次回忆，而非语言类的记忆，如复杂图形测验的回忆部分、符号数字模式测验的符号与数字匹配部分，都是不予提醒的？这并不清楚。提醒与不提醒的记忆是有区别的，提醒记忆（intentional memory task）更集中反映记忆能力；而非提醒记忆（incidental memory task）是附带的、伴随的记忆任务，有更多注意和（或）动机的成分。

大脑右半球损伤患者比左半球损伤患者的画图准确性差、扭曲更多、回忆成绩更差，但患者 CFT 的表现不能作为左右侧病灶判别的准确的预测指标。大脑后部（顶枕叶）病灶比前部（额叶）病灶患者的空间组织能力差。对该图的模仿有助于发现偏侧忽视。绝大部分遗忘是在模仿后几分钟出现的，正常人 3 min 和 30 min 延迟记忆的得分没有显著差别，这并不意味着间隔时间较长的延迟记忆可以忽略，因为不同病因的遗忘症患者有不同的记忆储存模式和能力。阿尔茨海默病（AD）患者早期即有 CFT 模仿和回忆能力受损。有迷路、无目的闲逛和不认识熟悉环境等症状的 AD 患者，其图画模仿能力也差。MMSE 的结构模仿仅 1 分，非语言的空间结构模仿和记忆能力在临床检查中常常被忽视，而视觉空间记忆缺损正是 AD 患者最重要的早期表现之一。我们在对偏侧忽视的研究中发现患者在完成 CFT 模仿图形（图 7-2-12）时采用的是从右到左渐渐推进的方式，而语义性痴呆患者相反，采用的是从左到右渐渐推进的方式，不同于上述定性分析的 6 种类型。

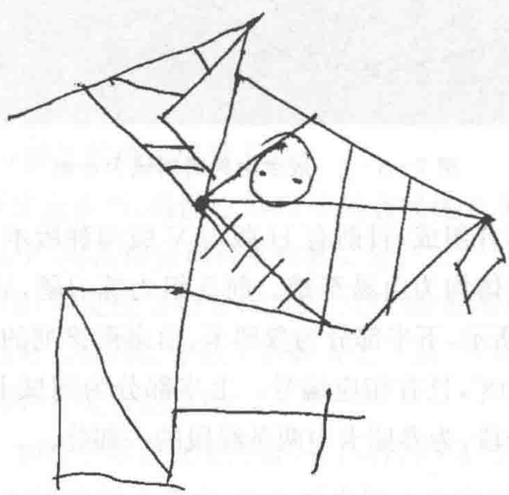


图 7-2-12　偏侧忽视患者画的 CFT 模仿图形

（郭起浩）

第三节 线方向判断测验(JLO)

线方向判断测验(judgement of line orientation, JLO)用于评估被试空间感知和定向能力,国外有针对7～96岁的常模。有关空间能力的测验繁多,1978年,Benton编制了JLO,最初的目的是在右利手大学生中研究大脑左右半球的不对称性,之后逐步演化为今天的版本,样例见图7-3-1。

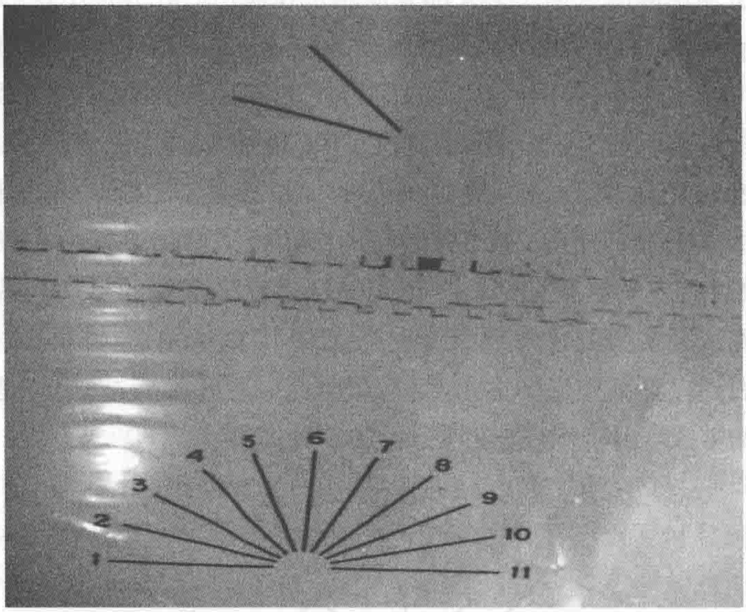


图7-3-1 线方向判断测试卡示例

JLO由35组卡片组成,目前有H版与V版两种版本,试题相同,仅呈现顺序有所差异,但总体均为由易至难。前5组为练习题,后30组为正式测试题。如图7-3-1所示,下半部分为参照卡,由扇形排列的11条线段组成,每两条线段夹角均为18°,且有相应编号。上半部分为测试卡,印刷有两条不同方向、不同长度的线段,为参照卡中两条线段的一部分。

一、实施方法

测试时,施测者面对受试者,取出测试卡片置于桌面,按照顺序依次翻开,呈45°角,使受试者能够清楚地看到测试卡及参照卡。要求受试者仔细观察测

试卡中的线段，指出其“在位置和所指方向上与参照卡中哪两条线段完全一致，告诉相应的编号”。如果受试者回答有困难，施测者可询问“请指出其中的一条线段，在位置和所指方向上与参照卡中的那一条线段完全一致，告诉我相应的编号”。指导语和练习卡可以重复，直至受试者能正确回答练习卡中的两条线段方可正式测试。否则，测试终止。答题无时间限制。整个测试耗时15～20 min。

二、评分

30 项标准版的得分为正确总数。如果要参照常模比较，需根据年龄和性别进行校正。50～64 岁受试者的总分在粗分基础上加 1 分；65～74 岁受试者加 3 分；不论何年龄段的女性受试者，均可额外再加 2 分。15 项简化版（O：奇数版；E：偶数版）的得分也为正确总数。如需参照常模评判，女性受试者需在粗分基础上加 1 分。

Ska 等曾将错误分为 4 种类型：① 同一象限错判（如将测试线段错判为同一象限的另一线段）；② 垂直或水平线段错判（例如，1、11、6 号线段的错误判断）；③ 不同象限错判（如将某一象限的线段错误判断为另一象限的线段）；④ 复合错误（水平/垂直线段错判及不同象限错判，即②+③）。

三、人口统计学影响因素

1. 年龄　测试得分随年龄呈曲线变化。儿童期随着年龄的增长测试得分提高，不同性别的儿童均至 13 岁达到成年人水平。无论是完整版还是简化版，50 岁以后的得分随年龄增长缓慢下降。

2. 性别　多项研究表明，男性在 JLO 中的表现优于女性，在 30 项标准版和 15 项简化版测试中，男性得分均较女性高 2 分。

3. 教育　Benton 等发现受教育程度对 JLO 的表现有中等程度的影响。标准版测试中，低教育人群的得分低于高教育的。对 15 项简化版的研究也表明，JLO 得分高与教育年限有相关性。

4. 智商（IQ）　相较于教育程度而言，受试者的 IQ 与 JLO 的相关程度更高，如果已将 IQ 影响因素纳入考虑，就无须再纳入教育程度。但在高智商群体中，这种相关性的强度略有减弱。Benton 等还曾报道，儿童的言语商（VIQ）与 JLO 得分无关。

5. 文化/种族　目前绝大多数有关 JLO 的报道都集中在欧美人群。

Benton 发现意大利和北美儿童在该测验中的表现相当；有关高加索人和非洲裔美国人的研究，则结果不一致。Trahan 报道两人种的卒中人群 JLO 的表现相似，但 Lucas 等则发现非洲裔美国老年人群 JLO 的得分较低。针对不同语言的研究相对较少，Rey 报道西班牙语人群 JLO 的中位得分为 24 分，划界分为 17 分，与 Benton 报道的英语人群（JLO 中位得分为 25 分，划界分为 15 分）的得分比较接近。我国近年才开始引进该测验，系统常模还未建立。复旦大学附属华山医院观察 180 例认知正常的老人发现，JLO 的中位得分为 22 分（未发表数据），略低于上述报道，可能是因为本组样本的平均年龄比较高（平均 70 岁）。

四、常模

西方 JLO 的最初常模数据来源于 1994 年 Benton 的工作，包括成年人和儿童。Benton 对 137 名 16～49 岁的成年人进行测试，发现男性平均得 26.6 分，女性平均为 23.3 分；93%的受试者的得分≥21 分，因此推荐 19～20 分为划界分。Benton 也对 221 名 7～14 岁的儿童进行测试，建立了儿童常模。

此后，Glamser 和 Turner 对 167 名 18～30 岁的大学生进行测试，女性平均得分为 22.80 分（标准差 5.66），男性平均为 26.21 分（标准差 4.37），与 Benton 的研究结果较接近。Mayo 诊所的 Ivnik 对高加索较大样本的人群进行了测试，将常模数据的年龄适用范围扩大到 55～97 岁，给出了各年龄段的百分位数，十分细致，有兴趣的读者可参阅表 7-3-1。除了粗分之外，Ivnik 还建立了标准分计算公式，将 JLO 粗分换算为校正年龄和教育程度的 MOANS 量表分。值得一提的是，Ivnik 最初的工作在高加索人群中开展，Lucas 在 2005 年应用参加 MOANS 研究的黑种人进行分析，发现其得分要低于白种人，提示了该测验的人种差异。目前我国还没有大规模的 JLO 常模数据，考虑到文化、语言、教育等多方因素的影响，读者在参考西方常模时应当谨慎。

对于简化版而言，1998 年 Woodard 等对 131 名 55～84 岁的高加索认知正常人群（MMSE 27.5±2.0）进行测试，采用的是 JLO 奇数版和偶数版（15 项，JLO-O、JLO-E），发现正常人群中极少得分低于 8 分。由此提出，JLO-O/E 得 10 分为临界，8～9 分提示轻度受损，6～7 分为中度受损，0～5 分为严重受损。同样，女性受试者得分加 1 分。

五、信度

已有多项研究验证了 JLO 标准版的折半信度较好，儿童中相关系数 $r=$

表 7-3-1　梅奥中心 56～79 岁 JLO(MOANS) 常模

百分位数	56～62(岁)	63～65(岁)	66～68(岁)	69～71(岁)	72～74(岁)	75～77(岁)	78～80(岁)	81～83(岁)	84～86(岁)	87～89(岁)	90～97(岁)	量表分
N	119	119	119	119	119	119	119	127	113	82	82	
<1	—	—	—	—	—	—	—	—	—	—	—	2
1	—	—	—	—	—	—	—	—	—	—	—	3
2	0～7	0～7	0～7	0～7	0～7	0～7	—	—	—	—	—	4
3～5	8～11	8～11	8～11	8～11	8～11	8～11	0～10	0～10	0～8	—	—	5
6～10	12～14	12～14	12～14	12～14	12～14	12～14	11～13	11～13	9～12	0～9	0～9	6
11～18	15～17	15～17	15～17	15～17	15～17	15～17	14～16	14～16	13～15	10～14	10～14	7
19～28	18～19	18～19	18～19	18～19	18～19	18～19	17～18	17～18	16～17	15～16	15～16	8
29～40	20	20	20	20	20	20	19～20	19	18	17～18	17～18	9
41～59	21～22	21～22	21～22	21～22	21～22	21～22	21～22	20～21	19～21	19～21	19～21	10
60～71	23	23	23	23	23	23	—	22	22	22	22	11
72～81	24	24	24	24	24	24	23～24	23～24	23～24	23	23	12
82～89	25～26	25～26	25～26	25～26	25～26	25～26	25	—	—	24	24	13
90～94	27～28	27～28	27～28	27～28	27～28	27～28	26	25～26	25	25	25	14
95～97	29	29	29	29	29	29	27～28	27～28	26～28	26～27	26～27	15
98	30	30	30	30	30	30	29	29	29	28～29	28～29	16
99	—	—	—	—	—	—	30	30	30	30	30	17
>99	—	—	—	—	—	—	—	—	—	—	—	18

MOANS 数据来源：梅奥中心老年人常模研究(Mayo's Older Normative Studies),Ivnik,1996.

0.84，成人中 $r=0.84\sim0.91$。但 15 项简化版(JLO-O, JLO-E)的信度系数相对低一些($r=0.69\sim0.75$)，与测验条目缩减有关。其他简化版，如 Winegarden(1998)报道的 20 项版的内部一致性为 0.80；而 Quall(2000)修订的 15 项版的内部一致性也达 0.82。

复测信度及学习效应：为避免学习效应，通常交替使用 JLO-V 版和 JLO-H 版，因为这两个版本的试题相同，仅呈现顺序有异，Woodard 认为这并不是严格意义上的替代版本。Benton 也曾对 37 名患者试用了上述两版，复测间隔 6 小时至 21 天不等，发现其信度系数为 0.90，并未发现学习效应。2001 年 Montse 给帕金森病患者和对照组进行测试(JLO-H)，20 min 后复测，也未发现学习效应。

六、效度

JLO 与韦氏智力测验的视空间分测验具有一定相关性，如积木测验($r=0.68$)、物品拼凑($r=0.69$)。与之相比，与韦氏智力测验中的语言类分测验则相关度较弱，如常识($r=0.45$)、词汇($r=0.28$)。

七、临床研究

凡是累及视空间能力的疾病状态，均可在 JLO 中有所体现，例如视觉忽略、阿尔茨海默病、Turner 综合征、帕金森病。值得注意的是，绝大部分研究分析的是 JLO 的正确总数，而较少关注错误类型。因此，阿尔茨海默病患者是否具有特征性的错误类型还未可知。Ska(1990)发现，与对照组相比，AD 患者更易发生不同象限错判。但 Finton(1998)并未发现 AD 患者的特征性错误。这种矛盾的研究结果可能与认知损害的不同严重程度有关。对于帕金森病，研究结果相对一致。Finton 和 Monste 都发现，PD 患者较正常对照组更易出现不同象限错判和水平线段错判，而较少出现同一象限错判。

Benton 等的经典研究提示，右侧大脑后部的损伤较左侧大脑半球损伤者的 JLO 表现更差。Trahan 和 Qualls 的研究也表明，右侧半球损伤，尤其是存在视觉忽略的患者，与左侧半球损伤的患者相比，JLO 测试的错误更多。Ng 则指出，尽管左侧顶叶病变也会影响 JLO 的表现，但右顶叶病损显然与更为严重的视觉功能障碍相关。上诉观点也得到了功能磁共振(fMRI)研究结果的支持。Gur 与 Ng 发现，右侧大脑半球后部在受试者完成该项测试的过程中占据主导地位，很可能是视觉信息处理的触发脑区。

抑郁和精神疾病(如精神分裂症)可影响测试表现,但非急性期精神分裂症患者的得分明显改善,与正常人基本相仿。

其他影响因素包括视力、测试时间和测试人群。

视力:由于该测验的视觉图像并不复杂,分辨率也较高,因此一般的视力下降、视物模糊并不影响测试成绩。当然,严重的眼部疾病(如黄斑变性)等仍会影响测验得分。

测试时间:选择哪个时间段进行测试对老年人的测验得分有一定的影响,但较轻微,目前并无统一的时段推荐。

测试人群:成年人的JLO测试表现不佳提示右侧半球病变,但儿童并非如此。Paquier(1999)发现,儿童JLO得分低并不一定提示脑病,也无法定位在某侧半球。本章节侧重于JLO在中老年人群中的应用,对儿童受试者并不重点阐述,但读者应了解,在儿童受试者中,该测验更多反映的是视觉-空间处理能力,而并非特异的脑部病变。

八、简化版

JLO有多种简化版,常用的为奇数版(O)和偶数版(E),分别为选取30项完整版(V版)中的奇数页或偶数页组成。Qualls等应用内部一致性分析,开发出15项的Q版和S版;Winegarden选取V版中的11~30项组成了20项的简化版。

JLO-E和JLO-O两版得分的相关度较高($r=0.71\sim0.81$),平均分和分值分布都比较接近。试题的呈现顺序对结果影响不大,提示,在对同一患者长期随访进行多次测试时可以选择不同顺序的替代版本。多项研究都证实,JLO-E/O与30项完整版的相关度很高($r=0.90\sim0.95$)。与完整版一样,简化版JLO-E/O也适用于多种神经、精神科疾病的患者群的测评,但能否鉴别左侧半球或右侧半球病变还未可知。

应当注意的是,不能将JLO-E/O的得分简单翻倍来对完整版常模做评判,应当参照简化版相应的常模。Qualls(2000)开发的JLO-Q同样也是15项,选择了JLO-V(标准版)中的16、9、6、2、12、30、7、17、19、28、20、21、26、24、22题组合而成(以上为Q版的施测顺序)。研究表明,JLO-Q与完整版得分高度相关($r=0.94$),识别功能损害与完整版也具有良好的一致性($\kappa=0.85$)。但在区分不同严重度(鉴别正常、轻/中/重度损害)的患者方面,JLO-Q与标准版判断的一致性差强人意。而在鉴别左侧半球和右侧半球卒中患者方面,无论是JLO-Q或是标准版,都没能获得理想的结果。此外,

Winegarden(1998)的20项简化版与标准版的相关度也较高(r=0.97)。同样,简化版需参照其相应的常模及划界分。

九、评价

JLO测验有许多优点。完成测验所需的肢体活动较少,也基本不需复杂的语言介导,评价了较基本的视觉空间能力。在北美和欧洲地区的正常值较接近,具有良好的信度,基本没有学习效应。

研究发现,年龄、性别、教育和智商对测验得分有一定的影响。55岁以下成年人的正常值考虑了年龄和性别的因素,但未校正教育或智商因素。55岁以上人群的常模分别校正了年龄+性别或年龄+教育/智商因素,但没有同时校正年龄、性别、教育/智商三个因素。该测验的"天花板效应"是其重要的局限性之一,许多受试者得到满分或接近满分。因此,仅凭JLO得分"正常"并不能断定视觉空间能力无损害;反之,得分低也并不一定有认知障碍,例如,针对美国黑种人的研究发现,JLO的低分可能与文化、测验熟悉程度、测试场所、受试者教育水平等均有关。

综上所述,JLO具有良好的信度、效度,是评价视空间能力的理想量表。然而,读者在使用时,有几个注意点:第一,对于老年受试者和认知功能严重受损的患者而言,JLO可能难度过高,受试者可能会在逐个测试条目上,如线段的斜率判定上,花费大量的时间。第二,尽管简化版与标准版有较高的相关性,但简化版并不能做认知损害严重度的分级判断。而且,简化版的重测信度数据还比较少。对于脑内偏侧病灶的定位、视知觉障碍的鉴别等,简化版是否适用,目前也没有充足的数据。因此,简化版可能更适用于筛选,或测试极易疲劳的状况;如需判定认知缺损严重程度,应当使用标准版。此外,将简化版分数翻倍后参阅标准版常模的做法是不可取的。最后,成年人中,JLO对右侧半球病变比较敏感。在儿童则主要用于评价视觉空间能力,尚不足以预测脑内病变或做病灶侧定位。

(赵倩华　郭起浩)

第四节　视觉物体与空间感知测验(VOSP)

视觉物体与空间感知测验(visual objective and space perception battery, VOSP)(Warrington 和 James, 1991)是用以评估受试者对物体和空间的感知能力的一组测验,适用于20～84岁的受试者,共包含8项分测验,由Warrington

等编制而成。该测验主要用于评价大脑右侧半球病变的认知损害。Warrington认为，物体感知和空间感知定位于不同脑区，在功能上相互独立，病变时可出现分离性损害。VOSP的8项分测验中第1～4项（字母辨识，剪影测验，物体识别，递进剪影测验）用于评估物体感知；而第5～8项（数点测验，位置判别，数字位置，立方体分析）用于评估空间感知能力。为确保受试者的视觉能力符合VOSP测试的要求，在正式测试之前，还设计了筛查测验。所有分测验均对大脑右侧半球病变较敏感，无须绘图或拼搭，对肢体操作和运动的要求极低。因此，肢体活动不便者，如卒中偏瘫、帕金森病患者等也适用。

一、实施方法

本节提供了各项分测验的操作说明和图示举例，读者可参考表7-4-1和图7-4-1。

表7-4-1　VOSP分测验描述

形状筛查测验（shape detection screening test）	由20张测试卡组成，如图7-4-1所示，要求受试者选择出其中隐含字母“X”的图形。记录正确数。筛查测验得分低于15分的受试者不适合接受VOSP正式测验
分测验1：字母辨识（incomplete letters）	如图7-4-1所示，向受试者呈现20个印刷不完整的英文字母图案（隐蔽30%～70%），要求受试者说出是什么字母。记录正确数，满分20分
分测验2：剪影测验（silhouettes）	本测验由15种动物和15种常用物品的图像组成，图像为从某个不常见视角观测的黑色剪影，测试卡由易至难呈现。记录正确数，满分30分
分测验3：物体识别（object decision）	每张测试卡上有4个剪影，其中之一为真实物体的剪影，另外三个是干扰项（非真实物体）。要求受试者选择出真实物体的剪影。记录正确数，满分20分
分测验4：递进剪影测验（progressive silhouettes）	两个物体（手枪和小号），以不同的视角呈现，逐步给出更为完整的剪影。每个物体从难至易逐一呈现，最多10次，直至受试者回答正确为止。记录每个物体受试者的尝试次数（最大为10+10=20分）
分测验5：数点测验（dot counting）	由10张测试卡组成，每张卡片印有随机排布的数个黑点。要求受试者准确答出黑点个数。记录回答正确的卡片数，满分10分
分测验6：位置判别（position discrimination）	由20张测试卡组成，每张卡片有两个方框，中间分别有一黑点。其中一个在方框的正中心，另一个黑点略微偏离正中一些。要求受试者选择黑点位于方框正中的那幅图。记录正确数，满分20分
分测验7：数字位置（number location）	由10张测试卡组成，每张卡片有上下排列的两个方框，上半部的方框中为随机排布的一些数字，下半部方框中有一个黑点。要求受试者对比两个方框后，说出上部方框中与下方黑点位置一致的相应数字。记录正确数，满分10分
分测验8：立方体分析（cube analysis）	由10张测试卡组成，每张卡片上有重叠摆放的数个立方体的侧面图，要求受试者说出立方体的个数（“隐藏”的立方体不计数）。记录回答正确的卡片数，满分10分

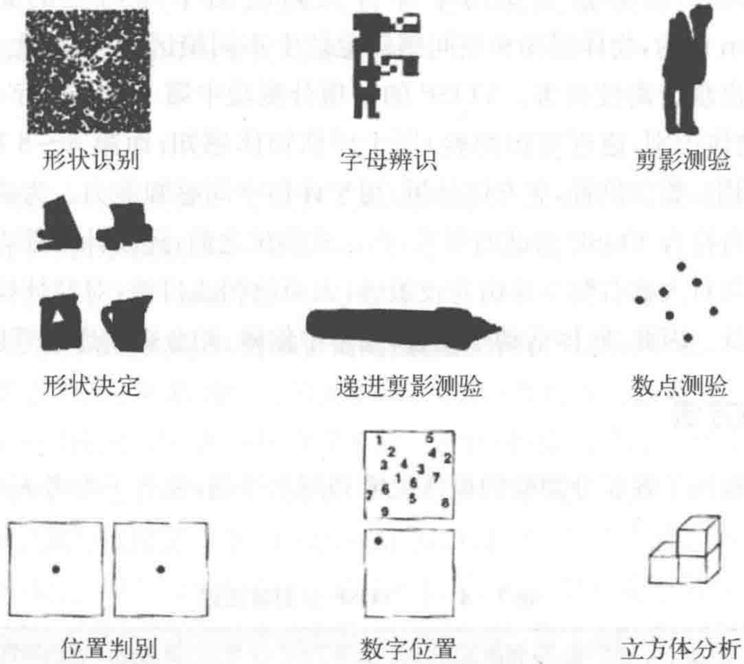


图 7-4-1　VOSP 分测验图示举例

每个分测验都有两次练习机会。各分测验测试顺序无特殊规定。测试不限时，每个分测验耗时 5～10 min，整个 VOSP 测试时间在 40～80 min。

二、评分

Warrington 的操作手册中附有判定异常损害的划界分（即得分低于相应年龄组认知正常受试者的 5%百分位数）；剪影测验和递进剪影测验，除 5%百分位数外，操作手册中还提供了 25%、50%、75%百分位数以供参考。

三、人口统计学影响因素

1. 年龄　研究表明，50 岁以后，受试者的所有分测验表现有所下降。部分测验表现（包括数字位置、物体识别、剪影测验、递进剪影测验）在受试者 70 岁后进一步下降；其余测验（包括：数点测验、字母辨识、位置判别）并未出现随年龄增长的下降趋势。Bonello 与 Herrera-Guzman 认为这可能与其“天花板效应”有关。但也有人认为，随着年龄增大，物体感知能力的衰退先于空间感知。

2. 性别　Benello（1997）认为该测验不受受试者性别的影响。但

Herrera-Guzman(2004)发现，8 项分测验中的 5 项(剪影、物体识别、递进剪影、位置辨别、立方体分析)测验，男女性的表现有所差异(表 7-4-2)。

表 7-4-2 西班牙语人群中 VOSP 各分测验的不同性别均分

VOSP 测验	男				女			
	n	*M*	*SD*	Min～Max	*n*	*M*	*SD*	Min～Max
字母辨识	45	18.96	1.15	15～20	45	19.33	0.90	17～20
剪影测验	45	21.04	3.42	11～28	45	18.44	4.65	8～27
物体识别	45	16.89	2.16	10～20	45	15.87	2.78	8～20
递进剪影测验	45	10.20	2.63	4～14	45	11.58	2.29	8～16
数点测验	45	9.80	0.40	9～10	45	9.71	0.59	8～10
位置判别	45	19.56	1.31	14～20	45	18.64	2.40	7～20
数字位置	45	8.51	2.03	1～10	45	8.60	1.50	4～10
立方体分析	45	9.00	1.71	1～10	45	8.13	1.98	2～10

(Herrera-Guzman, Clin Neuropsychocologist, 2004)

3. 教育/智商(IQ) 研究发现，以下分测验的表现与教育程度有较弱的相关性：数字位置、物体识别、剪影测验(表 7-4-3)。以下测验与智商 IQ 有较弱的相关性：剪影测验(r=0.29～0.30)、数点测验(r=0.27)、立方体分析(r=0.29)。

表 7-4-3 西班牙语人群中 VOSP 各分测验的不同教育程度均分

VOSP 测验	Basic				Medium				Higher			
	n	*M*	*SD*	Min～Max	*n*	*M*	*SD*	Min～Max	*n*	*M*	*SD*	Min～Max
字母辨识	30	18.87	1.11	16～20	30	19.10	1.18	15～20	30	19.47	0.73	18～20
剪影测验	30	18.97	3.69	10～26	30	19.03	4.87	26～28	30	21.23	3.87	13～28
物体识别	30	15.97	2.46	10～20	30	15.77	3.01	8～20	30	17.40	1.69	13～20
递进剪影测验	30	11.27	2.45	6～16	30	11.37	2.14	6～16	30	10.03	2.86	4～15
数点测验	30	9.73	0.45	9～10	30	9.67	0.61	8～10	30	9.87	0.43	8～10
位置判别	30	18.77	1.69	13～20	30	19.07	1.74	14～20	30	19.47	1.22	15～20
数字位置	30	8.13	1.77	4～10	30	8.57	1.98	4～10	30	8.97	1.52	4～10
立方体分析	30	8.23	1.61	5～10	30	8.30	1.56	5～10	30	9.17	1.12	6～10

注：教育分组：Basic，正规教育年限，0～5 年；Medium，6～11 年；Higher，12～20 年。
(Herrera-Guzman, Clin Neuropsychocologist, 2004)

4. 种族/文化 有关种族和文化背景是否会影响 VOSP 的表现，研究结果还有争议。目前有文献报道曾做过对照研究的是英国人群、美国人以及西班牙人。Herrera-Guzman(2004)比较了西班牙老年人与英国人和美国人的

VOSP 各分测验平均分，认为不存在显著差异(表 7-4-4)。Bonello(1997)报道了英国人群与美国人群在该组测验中的差异，但由于不同研究的受试者来源不同(医院病例与健康人群)，使得结果分析扑朔迷离。目前还没有中国人在该组测验中表现的报道，笔者的数据正在统计中。

表 7-4-4 不同人种的 VOSP 各分测验的比较

VOSP 测验	英国人群 (n=160)		美国人群 (n=111)		西班牙人群 (n=90)	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
字母辨识	18.8	1.4	19.46	0.73	19.14	1.04
剪影测验	22.2	4.0	20.40	3.77	19.74	4.46
物体识别	17.7	1.9	17.54	1.89	16.38	2.53
递进剪影测验	10.8	2.5	9.62	2.20	10.89	2.55
数点测验	9.9	0.2	9.77	0.61	9.76	0.5
位置判别	19.6	0.9	19.48	1.34	19.10	1.98
数字位置	9.4	1.1	9.08	1.31	8.56	1.77
立方体分析	9.2	1.2	9.54	0.8	8.57	1.89

(Herrera-Guzman, Clin Neuropsychocologist, 2004)

四、常模

VOSP 操作手册中提供的常模来自 Warrington 等在英国进行的两项独立的研究。其一为标准化研究，用以获得各个分测验不同年龄组人群的正常值；另一项为验证性研究，用以检验该测验对右侧大脑半球的敏感性，以及对左侧大脑半球病变的"非选择性"。标准化研究的 350 名受试者来自英国某医院，年龄 20～69 岁，均为非脑部病变的神经科病患(根据成人阅读测验 NART 估计，平均智商约 110)。考虑到年龄对测验结果的影响，操作手册中分别提供了 50 岁以下和 50 岁以上人群的测验均分和标准差。

然而，该常模不久之后便受到质疑。1997 年 Bonello 在美国密歇根对 111 名、年龄 50～84 岁、教育 4～18 年的认知正常老人(根据 Shipley 测验估计，平均智商约 112.5，与英国研究相近)进行测验。结果发现，如果采用操作手册中的划界分，不少认知正常的受试者会被误判为"认知损害"，在"剪影"和"数字位置"两项分测验中尤为严重。Bonello 还指出，操作手册中仅以 50 岁年龄为界提供常模可能并不合适，因为受试者在 70 岁前后测验的表现仍有显著差别。英、美这两项研究对象的来源、种族并不相同，造成结果差异的确切原因并不十分明了。我

国目前还没有 VOSP 的常模，读者在使用该测验，或参阅西方常模数据时，应尤其谨慎。表 7-4-5 为 Bonello 报道的美国人群的正常值数据。

表 7-4-5　美国人群 VOSP 各测验得分的正常值(均数与标准差)

分　测　验	<70 岁(n=52)	≥70 岁(n=59)
字母辨识	19.46(0.73)	19.12(1.37)
剪影测验	20.40(3.77)	15.53(3.87)
物体识别	17.54(1.89)	15.61(2.43)
递进剪影测验	9.62(2.20)	12.03(2.40)
数点测验	9.77(0.61)	9.68(0.68)
位置判别	19.48(1.34)	19.27(1.03)
数字位置	9.08(1.31)	7.12(2.88)
立方体分析	9.54(0.80)	8.54(2.05)

五、信度

1. 内部一致性　根据 Bonello 报道，对认知正常的老人，VOSP 的内部一致性总体而言不高(表 7-4-6)。各项分测验中，数字位置、剪影测验、立方体分析的内部一致性较高；但字母辨识、物体识别、递进剪影、数点测验的内部一致性相当低。递进剪影测验只设两种物体，可能是造成其内部一致性低的主要原因。而每项分测验中的各个测试卡的价值也不相同。Bonello 等甚至建议，如果对 VOSP 进行修订，原版中仅一部分测试卡(字母辨识中的 7/20 题、剪影测验中的 14/30 题、数点测验中的 2/10 题、位置辨别中的 9/20 题、数字位置中的 9/10 题、立方体分析中的 9/10 题)值得被保留下来，而递进剪影测验的两题均无保留价值。

表 7-4-6　VOSP 的内部一致性

相　关　系　数	分　　测　　验
很高(≥0.90)	—
较高(0.80～0.89)	数字位置
尚可(0.70～0.79)	立方体分析 剪影测验
临界(0.60～0.69)	位置判别
较低(≤0.59)	字母辨识 物体识别 递进剪影测验 数点测验

2. *重测信度及学习效应* Bird(2004)对99名39～75岁的健康成年受试者进行“剪影测验”，1个月后复测，重测信度较高($r=0.88$)，学习效应较弱，且不受年龄或智商的影响。

六、效度

Rapport等对Bonello研究的亚组人群进行因子分析，证实了VOSP存在物体感知和空间感知两个因子，印证了Warrington的“双因子”学说。但同时亦发现，VOSP中物体感知因子较强，其相关测验作为受试者结构能力检查手段更可靠。

七、临床研究

1991年，Warrington和James报道了一系列基于一侧脑部损害患者的研究。根据CT、血管造影或手术所见，将受试者分为左侧大脑半球或右侧大脑半球病变组。发现右侧大脑半球病变组的各项分测验表现均较左侧大脑半球病变组更差。在物体识别、递进剪影、数点测验、位置辨别、数字位置、立方体分析等分测验中，左侧大脑半球病变组与正常对照组的表现基本相近，提示该测验的假阳性率较低。

此外，Henry等发现，22号染色体长臂基因片段缺失导致的颚-心-面综合征患者出现VOSP表现的损害，与遗传因素导致的认知损害有关。Calderon发现，痴呆严重度匹配的路易体痴呆患者与阿尔茨海默病患者相比，VOSP多项分测验(字母辨识、物体识别、立方体分析)的表现更差。Sanchez-Pemaute则发现亨廷顿病患者的剪影测验与其脑内纹状体神经元细胞丢失有关。

八、评价

综上所述，VOSP测验证实了Warrington等有关物体感知和空间感知定位于独立的不同脑区并可选择性损害的理论。在正常人群中，部分VOSP测验的内部一致性较差，尚需在临床患者中进一步评价。剪影测验的重测信度较高，适合长期随访，监测患者的物体感知能力变化。对于其他分测验的复测研究还很缺乏，目前尚无法评价。

现已发现，最初英国研究的常模显然不适用于北美人群。其中的一个原因可能是，英国的研究选择院内患者作为研究对象，而不是社区正常人群。其二，标准化研究的350名受试者中，Warrington等对其中的200人进行了

VOSP 中 5 项分测验的测试，而对另外 150 人进行了其余 3 项分测验研究。双重样本的设计阻碍了对受试者疲劳度、测验熟悉程度等因素的分析。目前，已有北美人群、西班牙人群等的研究数据，是对 VOSP 测验常模的有益补充。

VOSP 的因子分析结果证实了 Warrington 的双因子理论。其中，物体感知因子显然更强，是反映受试者结构能力的更可靠指标。这也提示，对 VOSP 中的空间测验部分（尤其是位置辨别分测验）还需进一步改进。目前也已知，VOSP 能反映右侧大脑半球病损，但对于部分分测验（字母辨识、数点测验、位置辨别）的“天花板效应”是不容忽视的问题。因此，该测验对于轻度损害患者的识别能力可能受到限制。

（赵倩华　郭起浩）

第八章

执行功能的评估

执行功能的概念与脑结构基础见第一章第六节。常用的执行功能测验包括范畴测验(category test, CT)、认知估计测验(cognitive estimation test)、图案流畅性测验(design fluency test)、Ruff 图形流畅性测验(Ruff figural fluency test, RFFT)、威斯康星卡片分类测验(Wisconsin card sorting test, WCST)、迷宫测验、Raven 推理测验(Raven's progressive matrices, RPM)、Stroop 色词测验(SCWT)、连线测验(trail making test, TMT)、汉诺塔测验、韦氏智力测验的部分分测验(如相似性测验、图片排列测验)。最常用的执行功能测验是 WCST,但是该测验难度较高、耗时较长,一般用于精神分裂症、抑郁症等成年患者的检测,很少用于老年人认知受损的评估。本节介绍老人中更常用的 TMT、SCWT、BADS、WCST、CT 和 CST。

第一节 连线测验(TMT)

连线测验(TMT)是 1944 年美国陆军开发的测验,是 Halstead-Reitan 成套神经心理测验中的一个分测验,是目前世界上最普及、最常用的神经心理学测验之一,它反映注意、次序排列、心理灵活性、视觉搜索和运动功能,反映定势转移(set shifting)能力,同时反映手-眼协调能力、空间知觉和注意能力。定势转移是内源性注意控制机制,当两项任务竞争同一认知资源时,对两项任务相互转换进行控制的过程。TMT 分 A、B 两部分,其操作与提醒语言均有详细规定,简要描述是: TMT - A 部分(图 8 - 1 - 1),从 1～25 的数字按照顺序连起来。TMT - B 部分(图 8 - 1 - 2),按顺序连接,数字和字母交替。正式开始之前均有练习。

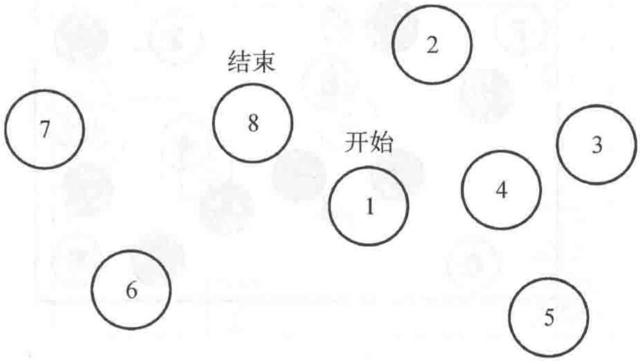


图 8-1-1　TMT-A 练习题

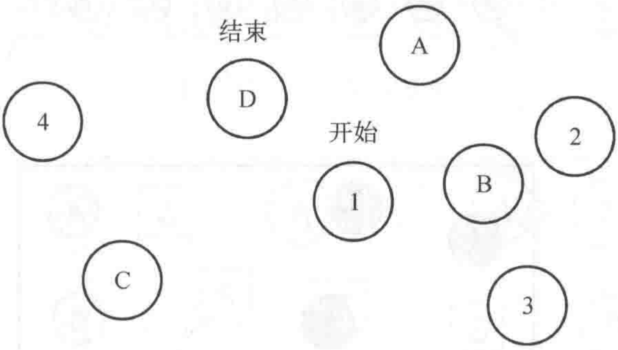


图 8-1-2　TMT-B 练习题

一、版本介绍

除了经典的 TMT 版本，还有许多变异版本。

1. **口头 TMT**　口头 TMT 省略了视觉运动成分，适合视觉障碍和利手瘫痪的受试者，要求 1-A，2-B，3-C……这样数字-字母配对，到 13-M 止。

2. **着色 TMT**　由于 TMT-B 部分是数字和字母交替连接，而非西方文化背景的人群对字母次序不熟悉，最常见的跨文化版本是着色 TMT，又称为 CTT(color trails test)(图 8-1-3)，有圆圈底色为红色与黄色的，也有用灰色与白色的，用两种不同颜色的数字代替数字和字母，体现文化公平，但在不同种族的(非裔、西班牙裔和高加索裔)美国人中依然有显著差异。

3. **形状连线测验(shape trails test，STT)**　复旦大学附属华山医院神经心理室采用的 TMT 是将数字包含在正方形和圆形两种图形中(图 8-1-4、8-1-5)。TMT-B 部分，按顺序连接数字时两种图形要交替进行。这种基于不同形状的 TMT，我们称为 STT。

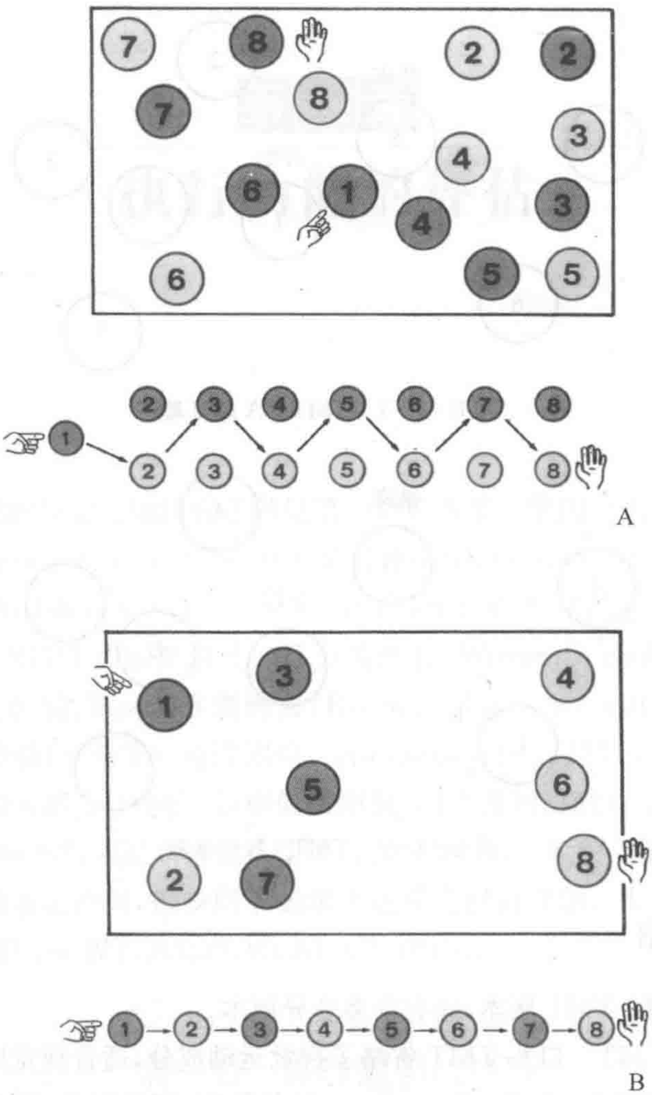


图 8-1-3　CTT 练习题

4. 综合 TMT（comprehensive trail making test，CTMT）　Reynolds 2002 年开发的一套 TMT，称为 CTMT，共分 5 个部分。第一部分与 TMT-A 相似；第二部分是在第一部分的基础上增加了 29 个空心圆圈作为分心；第三部分是在第一部分的基础上增加了 2 种类型的分心，如 13 个空心圆圈、19 个圆圈内有不规则图形的圆圈；第四部分为阿拉伯数字（如 1，7，在圆圈内）与拼写数字（如 six，nine，在长方形内）之间转移连接；第五部分是数字与字母交替连接，类似 TMT-B 部分，但是有 50 个空心圆圈作为分心。该版本目前没有中文版。

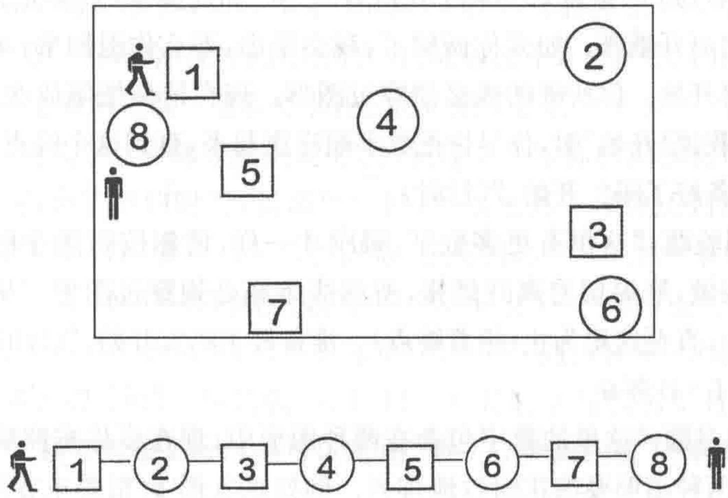


图 8-1-4　STT 练习题 1

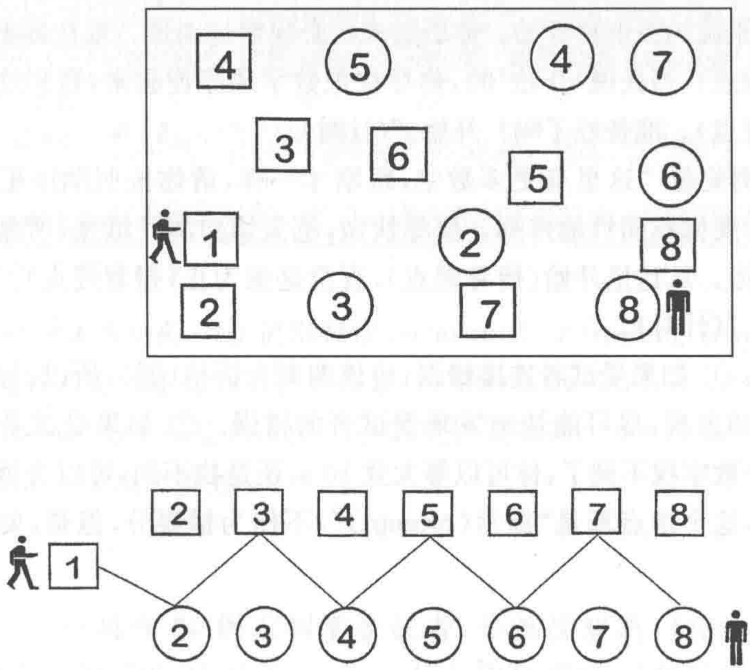


图 8-1-5　STT 练习题 2

二、指导语

1. STT-A 部分

(1) 练习题:“这里有些数字,当我说‘开始’时,请你把它们顺序连起来。

从1(指着1),到2(指着2),再到3(指着3)……如此类推,直至完成为止。注意笔尖切勿离开纸张。如果你做错了,我会指出,那么你退回来,从前一个正确的数字再开始。你所画的线必须穿过图形。现在请你把笔放在这里(指着起点)。当我说'开始'时,你尽快把数字顺序连起来,直到这个终点为止(指着终点)。准备好了吗?开始。"(计时)。

(2) 测验题:"这里有更多数字,跟刚才一样,请你按照顺序把它们连起来。要尽快做,笔尖切勿离开纸张,所画线条亦必须穿过图形。从这里开始(指着起点),直至这里为止(指着终点)。准备好了吗?开始。"(计时)。

2. STT-B部分

(1) 练习题:"这里的数字包含在两种图形中,现在要你按照顺序把数字连起来,而两种图形要梅花间竹地排列。例如正方形1(指着正方形1),到圆形2(指着圆形2),再到正方形3(指着正方形3)…… 如此类推,直到终点。注意:形状是要梅花间竹地交替排列的。如果你做错,我会指出,那你便要从最前面那个正确的图形再开始。你所画的线必须穿过图形。现在你把笔放在这里(指着起点),当我说'开始'时,你尽快把数字顺序连起来,直到这个终点为止(指着终点)。准备好了吗?开始。"(计时)。

(2) 测验题:"这里有更多数字,跟刚才一样,请你按照顺序把它们连起来,而图形要梅花间竹地排列。要尽快做,笔尖切勿离开纸张,所画线条亦必须穿过图形。从这里开始(指着起点),直至这里为止(指着终点)。准备好了吗?开始。"(计时)。

注意:① 如果受试者连接错误,应该即刻告诉他(她),所以,你要密切观察受试者的表现,尽可能快地发现受试者的错误。② 如果受试者手遮住答案,下一个数字找不到了,你可以等大致10 s,还是找不到,可以告诉受试者数字在哪里,这个指点算是"提示(prompt)",不作为错误分,但是,要记录提示次数。

三、评分

1. *评分方法* 让受试者开始连接到结束,记录总耗时数:① TMT-A和TMT-B的耗时数,练习与测试题,共4个指标。② 推导指标:差值(difference score)=B耗时数-A耗时数;比率值(ratio score)= B/A耗时数;对数(Log)=Log B∶A;常比(proportional score)=B-A/A,因为与B/A直线相关,所以不主张采用。③ 连接错误次数。④ 提示次数。

2. 评分时间　给予 4 min 界定，4 min 内完成，得满分 25 分，没有完成者，连接到那个数字得该分；推导指标、错误数、提示数等分析指标同上。

3. 第一分钟到达数　笔者要求评定员记录受试者完成 STT－B 部分时第一分钟达到的数字与全部完成的耗时数，发现 2 个指标的识别力没有显著差异，也就是说，为了节约时间，提高依从性，可以把第一分钟达到的数字作为 STT－B 的分析指标（表 8－1－1）。但是，在美国，55～69 岁中学以上教育程度的正常老人的 TMT－B 的耗时数为 60～70 s（表 8－1－2），也就是接近一半的受试者耗时不到 1 min，所以，并不需要“第一分钟到达数字”指标。但是，在日本，正常老人的 TMT－日文版本（TMT－B 是数字与假名交替）的耗时数（表 8－1－3）通常大于 2 min，是可以采用“第一分钟到达数”指标以节约测验时间的。

表 8－1－1　根据 636 例正常老人制定的 STT 的划界分

指　标	$\bar{\chi}-1.5SD$			$\bar{\chi}-1SD$		
	50～59 岁	60～69 岁	70～79 岁	50～59 岁	60～69 岁	70～79 岁
STT－A 中学教育	85	90	120	70	80	90
STT－A 大学教育	70	80	85	65	67	75
STT－B 中学教育	200	230	290	180	200	260
STT－B 大学教育	200	210	240	180	190	220

资料来源：笔者即将发表的论文。

4. 连接错误的分类　① 追踪错误（sequential or tracking errors）：指连接到错误的数字或字母上。② 持续错误（perseverative errors）：在 TMT－B 部分，不能在数字与字母之间转换。③ 邻近错误（proximity errors）：连接到最近但错误的数字或字母上。

四、评价

Lee（2000）调查 84 例右利手受试者，分英汉双语（Chinese-English bilingual，CEB）和英语（English monolingual，EM）两组，结果发现 EM 组的 TMT 和 CTT 高度相关，而 CEB 组两者没有显著相关性，除了 TMT－A，TMT－B、CTT－A 和 CTT－B 这 3 部分在 2 组之间没有显著差异。Dugbartey 认为 CTT－B 和 TMT－B 没有等效性。CTT 中可能存在 Stroop 效应。

TMT－B 反映的是快速视觉搜索、视觉空间排序和认知定势转移。TMT

表 8 - 1 - 2　美国正常老人的 TMT 表现

年龄(岁)	教育(年)	例数	Part A	Part B	教育(年)	例数	Part A	Part B
55～59	0～12	58	35.1(10.94)	78.84(19.09)	>12	37	31.72(10.14)	68.74(21.02)
60～64	0～12	55	33.22(9.1)	74.55(19.55)	>12	31	31.32(6.96)	64.58(18.59)
65～69	0～12	65	39.14(11.84)	91.32(28.89)	>12	32	33.84(6.69)	67.12(9.31)
70～74	0～12	76	42.47(15.15)	109.95(35.15)	>12	30	40.13(14.48)	86.27(24.07)
75～79	0～12	74	50.81(17.44)	130.61(45.74)	>12	34	41.74(15.32)	100.68(44.16)
80～84	0～12	84	58.19(23.31)	152.74(65.68)	>12	34	55.32(21.38)	132.15(42.95)
85～89	0～12	29	57.56(21.54)	167.69(78.5)	>12	13	63.46(29.22)	140.54(75.38)

资料来源：Tombaugh TN. Trail making test A and B: normative data stratified by age and education. Archives of Clinical Neuropsychology, 19(2004) 203 - 214.

表 8 - 1 - 3　日本正常老人的 TMT - 日文版本的表现

年龄(岁)	教育(年)	例数	Part A	Part B	教育(年)	例数	Part A	Part B
65～69	0～9	11	53.4(19.1)	140.6(60.8)	>10	13	40.0(9.0)	105.0(22.7)
70～74	0～9	10	45.2(15.1)	128.4(62.2)	>10	17	49.8(19.0)	116.7(55.9)
75～79	0～9	18	57.4(11.9)	152.8(59.5)	>10	10	51.4(12.6)	135.2(57.5)
80～85	0～9	8	60.6(18.0)	176.0(70.1)	>10	5	79.6(19.5)	174.0(54.2)

资料来源：原田浩美. 高级脑机能研究. 2006,26(1)：16 - 21.

对闭合性脑损伤、酒精中毒、药物依赖高度敏感，但对甲醛暴露受害者不敏感。Cahn 评估 238 例正常老人、77 例 AD 危险人群和 45 例 AD 患者，TMT - A 的平均得分分别为 48 s、56 s 和 84 s，TMT - B 的平均得分分别为 124 s、173 s 和 228 s，能有效区分这 3 组样本，TMT - B 的敏感性（87%）优于 TMT - A（69%），特异性相似（分别为 90%和 88%）。TMT 是识别轻度痴呆最敏感、最常用的评估工具之一。严重度匹配的 AD 和血管性痴呆之间没有显著差异。以往曾经将 TMT - A 和 TMT - B 的表现定位在左额叶，但近来发现左额叶损伤和右额叶损伤患者的 TMT 耗时数并没有显著差异，病灶容积与测验得分也没有相关性。

第二节 Stroop 色词测验(SCWT)

Stroop 色词测验（Stroop's color word test, SCWT）的目的是通过测量认知控制来评估受试者保持心中目标，抑制一个习惯性反应，而倾向一个较不熟悉的反应的能力，评估选择性注意及认知灵活性。

一、Stroop 色词测验的不同版本

Stroop 于 1935 年开发了 SCWT，其范例可追溯至 Cattell 于 19 世纪末的工作。目前发展出 10 多个版本的 Stroop 色词测验，主要区别在于使用卡片的数目（卡片的数量从 2 张到 4 张不等）、每张卡片的长度（即字数，少则 17 个、多则 176 个，如 Victoria 版本采用 24 个，D - K 执行功能系统版本采用 50 个，Golden 和 Comalli 版本采用 100 个）、字的颜色（少则 3 种、多则 5 种，如 Comalli 版本、Delis 版本和 Graf 版本为 3 种；Victoria 及 Golden 版本为 4 种；而原始的 Stroop 版本则为 5 种）以及评分方法（有的采用完成一定字数的时间消耗，有的采用限定时间内的完成字数）。举例来说，有时印有颜色名称的黑色字体卡片（卡片 A）会被省略（如 Trenerry 版本，1989），有时也会加入一张字体颜色与颜色名称一致的卡片（如 Graf 版本，1995）。有些测验不是单词，而是使用彩色的斑块、圆点（如 Comalli 版本、Victoria 版本、D - K 执行功能系统）或符号“X”（Golden 版本，1976；Graf 版本，1995）。D - K 执行功能系统还涵盖第 4 项条件，要求受试者需在命名字体颜色与命名与其不符的颜色文字间相互交错。这些不同版本的局限性包括：缺乏正常受试者错误得分的数

据，未能校正干扰试验中的普遍性减慢（generalized slowing），缺乏详细的年龄校正的正常受试者的数据，以及缺乏儿童中应用的数据。

1. 经典版　经典的CWT由四部分组成。第一部分（卡片A），受试者读出黑色印刷的表示颜色的字（蓝，绿，红，褐，紫）。第二部分（卡片B），受试者读出不同颜色（蓝，绿，红，黄）印刷的字（蓝，绿，红，褐，紫），要求忽略字的颜色（字与印刷它所用的颜色相矛盾，如“蓝”用绿色印刷）。第三部分（卡片C），受试者需说出彩色方块的颜色（蓝，绿，红，褐，紫）。第四部分，再次给予受试者卡片B，但受试者需说出字的印刷颜色而不是读出字。此试验主要观察受试者对于颜色名称用不匹配的颜色呈现时的表现。Stroop报道正常人读出字体颜色不同的颜色名称的速度与使用黑色字体时一样快（第二部分对比第一部分）。然而，当要求受试者读出字体颜色而非文字本身时（第四部分对比第三部分），完成任务的时间则明显延长。对于颜色命名速度的减退称为“颜色命名干扰效应”。

2. Victoria版　Victoria版本（Victoria Stroop test，VST）（Regard，1981）与Perret（1974）设计的版本相似并且有许多优势。首先，它很简短。与其他测验相比在每个部分都有大量的项目（如60～112个单词），而VST 3个部分（命名圆点的颜色，文字本身，以及以不对应颜色印刷的文字）各自只有24个项目。有证据显示较短的测验时间可能对于识别对这项任务有困难的个体更可取。Klein等（1997）发现在100项的Stroop版本中，老年人较年轻的成人受到的干扰更多，并且这个效应在前40项比后60项更突出。从而，VST不仅测试时间短，而且可能对于检测反应抑制的困难程度较理想，因为避免了受试者任务延长导致的学习效应。第二，可评估与信息加工速度相对独立的分数，包括可校正普遍性减慢的错误分数和干扰分数。第三，可获得一个合理的正常人的数据库（常模）。最后，VST是公开的，使用者可制作自己的测验卡片。

VST包括3张21.5 cm×14 cm的卡片：圆点（D部分）、文字（W部分）及颜色（C部分）。每张包含6行，每行有4个项目（Helvetica字体，28分）。每行中间距离1 cm。在D（圆点）部分中，受试者必须以最快的速度读出蓝、绿、红或黄颜色的24个圆点。每个颜色出现6次，并且4个颜色以一个伪随机的顺序排列在阵列内，每个颜色在每一行都出现一次。与Stroop色词测验不同，VST的W（文字）部分与D部分相似，只是圆点被一些常见的文字取代（when，hard，over），以小写字母打印。要求受试者说出字体颜色而忽略它的

含义。C(颜色)部分与D部分及W部分相似，但彩色的文字为小写字母打印的颜色名称“蓝色，绿色，红色和黄色”，且字体颜色与颜色名称矛盾(如“红色”以蓝色字体打印)。后面这项任务要求受试者抑制原本的自动阅读反应而产生一个更费力的命名反应。干扰效应决定于干扰任务比控制任务中命名颜色所需要多少额外的时间。

3. Golden版 Golden版本经常被使用。它包括一页黑色字体的100个颜色文字(红，绿，蓝)，一页使用红、蓝或绿色字体的100个“X”，及一页采用第一页的100个文字(红，绿，蓝)和第二页的字体颜色(字体颜色与文字不匹配)。受试者阅读每一页，以最快的速度一栏一栏往下读出文字或说出字体颜色，限时45 s。此测试基于每一页完成的项目个数一共产生3个分数。此外，也可以计算出干扰分数。

4. Comalli版 此版本包括3张24 cm×24 cm的白色卡片，各自包含100个项目(10×10)。第一张卡片包含黑色字体的文字“红色、蓝色和绿色”，随机排列成10×10的矩阵。第二张卡片包含红、蓝、绿色的斑块，随机排列成10×10的矩阵。第三张卡片包含以不匹配字体颜色打印的颜色名称。完成每张卡片所需要的时间作为一个独立的变量。Comalli-Kaplan修改版中要求先呈现颜色命名卡片且需将错误做记录。做此修改的理由为：① 色盲的个体可以被快速识别；② 此过程通过在干扰测试之前展示文字阅读部分从而将干扰效应最大化。Comalli版本有西班牙文版本，用于6～12岁的儿童。D-K执行功能系统采用Comalli版的修改版。

5. 华山版 根据现有版本的使用资料与我们自己的使用经验，笔者编制了华山版Stroop色词测验，它有如下特点：① 难度中等：共3张卡片，每张50个项目，卡片A是阅读黑色印刷的汉字(红、黄、蓝、绿)，卡片B是4种颜色的圆点，卡片C是文字颜色与文字意义不一致的汉字；② 记录耗时更客观：增加了卡片C的练习题，只有练习题阅读正确才开始做卡片C；正式开始后记录耗时，出现的阅读错误不予提醒(但予记录，便于分析错误类型)。由于不用针对每个错误进行提醒，受试者的读速没有被干扰，所以耗时数的记录应该更客观。

二、操作

1. Victoria版 在这个版本中，3张卡片依照固定的顺序呈现：圆点(D部分)，文字(W部分)及颜色(C部分)。指导语如下。D部分：“以最快的速度

说出圆点的颜色。从这里开始并沿着每行由左至右。"指示受试者的眼睛沿着每行由左至右。W 部分:"这次,以最快的速度说出文字的颜色。从这里开始并沿着每行由左至右。"如果必要可特别澄清:"说出文字的字体颜色。"C 部分:"再次以最快的速度说出文字的颜色。"必要时可澄清:"不要念出文字,请告诉我文字的字体颜色。"每个部分文字命名的错误若受试者没有自发更正,则立即由测试者更正,接着指示受试者以最快的速度继续。测试者在给完操作指示后立即启动计时器,需记录每个部分中产生的错误数及所花费的时间。

2. Golden 版　要求受试者从最左上方的文字开始,顺着往下读。在颜色命名测验中,要求受试者说出项目的颜色,其余要求皆与先前相同。在颜色-文字干扰试验中,要求受试者沿着每栏往下读,命名文字的颜色而忽略文字本身的含义。在所有试验中,皆要求受试者以最快的速度完成。错误由测试者单独指出,纠正受试者并继续。45 s 后,记录每个测试中最后说出的项目。不记录错误数目。

三、评分

1. Victoria 版　在每个部分,由测试者记录完成所需的时间及错误数量。受试者自发更正的也算正确。研究者一般采用"差异分"指标:定义为干扰测试(C 部分)与颜色部分(D 部分)所需的时间差。

Victoria 版的常模性资料见表 8-2-1。使用重叠的年龄组,包含 272 位健康成年人,年龄在 18～94 岁之间,居住于维多利亚、哥伦比亚或多伦多,64%为女性,平均受教育年限 13 年。他们通过大学老年数据库、老年中心、在社区及大学张贴广告招募。参与者经过筛查排除可能影响认知功能的神经系统疾病(失去意识 1 h 以上、卒中、癫痫、多发性硬化)及心理障碍(需住院的抑郁或焦虑)。所有受试者都能说流利的英语。

表 8-2-1　不同年龄组 Victoria 版耗时数与错误数表现(时间单位: s)

年龄(岁)	18～39	30～49	40～59	50～64	60～69
D 部分	11.0(2.5)	11.1(1.9)	12.3(2.4)	12.0(2.3)	12.1(2.3)
W 部分	13.0(2.9)	13.9(2.6)	15.2(2.9)	15.4(3.2)	15.9(5.1)
C 部分	22.1(7.2)	25.7(9.0)	27.8(8.2)	28.5(9.5)	29.4(9.0)
错误数	0.8(1.0)	0.8(1.0)	0.7(1.0)	0.6(1.0)	0.5(0.9)

（续表）

年龄（岁）	65～74	70～79	75～84	80～94
D部分	13.3(3.6)	14.2(3.9)	15.1(3.8)	15.9(4.2)
W部分	16.9(5.1)	18.6(5.4)	20.7(6.7)	22.1(6.0)
C部分	32.6(9.6)	37.1(11.9)	43.3(17.7)	50.4(23.9)
错误数	0.6(1.2)	1.1(1.6)	1.7(1.8)	2.1(2.0)

资料来源：Strauss E, Sherman E, Spreen O. A compendium of neuropsychological tests.（Third edition）. Oxford university press，2006：487.

2. Golden 版　这个版本产生 3 个分数：文字阅读分数（W）为 45 s 内完成的项目个数；颜色命名分数（C）为 45 s 内完成的项目个数；而颜色-文字分数（CW）为 45 s 内完成的项目个数。另外还可计算干扰分数。

Golden 版评分系统包括根据正常个体年龄和教育年限表现的预测值做出的调整分数。测验手册的表Ⅰ～Ⅲ提供 W、C 和 CW 的预测值。每项评分都由原始分数减去年龄/教育程度调整分数，所得数值再根据手册的表格Ⅳ转化为 T 分数（M＝50，SD＝10）。CW 减去 CW 的预测值可得出 15 岁以上受试者的干扰 T 分数（根据手册中的表格Ⅴ和Ⅵ得出）。对于年龄 5～14 岁的受试者，干扰 T 分数则是颜色-文字 T 分数与颜色 T 分数之间的差值。W、C 和 CW 的 T 分数，分值越高则说明表现越好。干扰分数较低的（T 分数＜40 分）则提示存在问题。已经有研究者对其样本量大小、干扰分数计算方法等提出质疑。

四、相关因素

1. 性别　虽然女性似乎有较杰出的颜色命名技巧，但在颜色-文字干扰部分中并未呈现优势或优势非常小。

2. 教育/IQ　教育程度与成人 Stroop 干扰分数中等相关（＜0.30）。在非裔美籍人群中，Moering（2004）发现教育程度对 Stroop 分数的影响最大，合计有 8%～26%的差别，其次是性别与年龄。但 Lucas（2005）报道无论教育程度还是年龄，对干扰试验表现的差异都为 8%～9%。与教育程度相比，IQ 与测试分数间表现出更强的关联性。随着测试的难度增高，受教育年数的相关性随之下降，但智商的相关性则随之上升。总的来说，个体的智力越高，他们受到的干扰则越少。

3. 种族或语言　种族对测试的结果也存在影响，纵然将教育程度也考虑在内，非裔美籍族群的表现皆比白种人要差。双语人群（西班牙文-英文）的表现也较单语种人群表现差，尤其是颜色命名部分。

4. 视力与阅读　视觉能力非常重要。在老年人中，结果的混淆可能与年

龄所致的视觉敏锐度或色觉的下降相关。Van Boxtel(2001)检验52～84岁的中老年人时发现经过年龄、性别、教育程度的调整，较低的敏感性与文字命名所需时间较长相关，色盲与较慢的颜色命名相关，而距离敏锐度的下降与干扰卡片的速度下降相关。Stroop色词测验表现中，一半由年龄解释的差距也可由视觉功能的不同来解释。显然，色盲应避免使用此试验；但这些发现提示，即使是正常衰老过程中视觉功能的微小缺陷都可影响Stroop测验的表现。

干扰的程度同样也依赖于受试者对刺激物的熟悉度，以及对材料语义方面的理解。阅读时，自动化反应的程度也是一项重要的因素。

5. 等效版本　虽然Stroop(1935)以卡片上的项目顺序倒过来读做了一个形式相当的测验版本，大多数测试者还是以同样的一套卡片重新测试受试者。Sachs发展了5种与Dodrill版本形式相似的测试。平行版本的可靠性系数为0.82。基于明显练习效应的现状，希望在相同或相似的版本上通过重复测验记录差异的测试者，应保证受试者在正式测验前有足够的练习。

五、效度

1. 测试间的相关性　在正常人群中，Victoria版测试间的相关性中等，而Golden版则是中等偏上，提示他们相似但不完全相同。

2. 与其他测验的相关性　干扰分数与其他注意力测试的方法之间有良好的相关性，包括连续操作任务(CPT)和连续加法测验(PASAT)中的遗漏错误。干扰分数与其他优势反应抑制(如连线测验A和B间的差异得分)呈中等相关(r=0.55)。

3. 工作记忆对Stroop干扰有贡献　根据Kane和Engle(2003)的报道，工作记忆能力的个体差异可预测Stroop测验的表现，提示在面对来自习惯的竞争时，维持目标的重要性。这些作者提出Stroop效应的双重机制观点，也就是Stroop干扰可反映受试者对于维持"忽略文字内涵"这项目标的失败。干扰也可能反映在成功激活的目标中，由于反应竞争导致耗时延长。

4. Stroop干扰也可能反映受试者的概括能力和处理速度　正常老年人在Stroop词色与词义不一致的卡片中，85%的年龄相关差异取决于总体的处理速度。在临床样本中，不像威斯康星卡片分类测验(WCST)和连线测验B部分这类测查执行功能的测试，Stroop干扰承载一些似乎能代表处理速度的因素。此外，Stroop干扰还涉及语义系统和计划性。Hanes(1996)发现在精神分裂症、帕金森病和亨廷顿病的患者中，Stroop干扰与语义流畅性任务的表现(r=0.58)和伦敦塔测验(r=0.65)有很强的相关性，但与其他的任务只有很小的相关性，如

Rey复杂图形测验的延迟回忆(r=0.31)和钉板测验得分(r=0.12)。

六、临床研究

Stroop干扰效应出现在精神分裂症、帕金森病、亨廷顿病、弗里德赖希共济失调、MCI与AD、孕期酒精暴露、慢性酗酒、HIV感染、儿童良性局灶性癫痫和儿童多动症等疾病中。在轻度的颅外伤患者中使用一个更具挑战的Stroop版本可以更容易地引出干扰效应。这个更具挑战的版本主要是在颜色文字分测验中进行修改——在1/5随机选取的项目周围画上长方形线条。在框住的项目中,要求受试者说出文字而非字体颜色。任务对于不同项目阅读和命名间注意力的转换要求较高,从而提升了任务的复杂程度。

Stroop干扰效应也与痴呆相关。抑制程序的降低常于AD早期即出现,而且干扰效应的强度随着AD严重程度的增加而增加。但Stroop表现与AD严重程度只是中等相关(相关系数约0.30),提示与AD早期识别的比较,此任务对于监测疾病的进展可能作用较小。干扰的增加在无痴呆皮质下腔隙性脑梗死患者中也有报道,白质高信号的程度与Stroop表现相关。笔者比较正常老人、MCI患者和AD患者发现,轻度AD患者在处理速度和正确性的关系方面,以牺牲正确数来换取阅读速度,而MCI组的应答策略是试图延长阅读时间(减慢阅读速度)来换取阅读正确性。

前额叶皮质似乎对年龄增长特别敏感,因为这个区域是影响Stroop表现的重要因素,随着年龄增长干扰效应增高。Uttl和Graf(1997)统计了一个年龄跨度较大的样本(12~83岁),他们发现年龄在"文字颜色与意义"不一致情况的结果中有一定的影响。他们假设Stroop干扰的年龄效应是由于处理速度随着年龄下降,而并非一项特定的认知功能降低。但是后来的研究发现,就算将普遍性的减慢列入协变量,年龄仍与干扰的增加有显著相关性,提示Stroop的年龄相关的得分下降与认知处理相关,而并非普遍性降低引起。Shilling(2002)建议,智力的年龄相关递减或许可解释Stroop的年龄相关改变。

Stroop色词测验表现的降低也见于抑郁患者,尤其是患双相障碍的患者较单纯抑郁和焦虑的患者尤甚。这意味着测试者不应完全将表现不佳归因于症状本身。

七、评价

神经影像和电生理研究显示额叶是活动最持续的部位。与这些相符,局

灶性额叶病变的患者在 Stroop 测验中表现出超出正常范围的干扰。一项近期的 meta 分析比较额叶损害与后脑损害患者，在 Stroop 的所有测验中皆发现显著差异。但单独的 Stroop 表现对于分辨额叶或非额叶组并不显著，组间少部分病例的分离从而导致敏感性及特异性较差。额叶的多个区域被认为对于 Stroop 测验的认知控制起重要作用，包括外侧前额叶皮质和前扣带回皮质。基于对健康个体的 fMRI 研究，Kerns(2004)认为前额叶皮质与实施控制以克服冲突的过程相关，而前扣带回皮质(ACC)可以识别行动计划间的冲突，并实现冲突监测的功能。虽然额叶功能显得非常重要，但 Stroop 测验任务是借由一个基础更为广泛的系统介导的。Stroop 表现与 AD 患者海马和数个新皮质区域包括后脑区域的病理表现相关。另外，健康成年人在实行 Stroop 测验过程中的 fMRI 活动不仅是额叶区域，也包含了下颞叶、顶叶皮质以及尾状核。

根据 Van Gorp(1999)的报道，虽然 Stroop 测试本身并不能识别诈病患者，但诈病患者会花费更多时间完成颜色命名和干扰测试。Lu(2004)提出 Stroop 测试对主诉自己是完全文盲的诈病患者特别有用。他们描述了 6 位患者主诉其无法进行文字阅读试验，但在颜色-文字干扰试验中，他们都犯了读出文字的错误。6 位中的 5 位在颜色-文字相关的干扰试验中，完成速度大幅减慢，提示他们其实在抑制阅读反应。主诉完全没有阅读能力的情况很少见，然而，当有这种情况时，抑制 Stroop 效应的失败可作为鉴别此症状表现的有效工具。

Stroop 色词测验是用于检测注意力和反应抑制最老也是最广泛应用的工具之一。目前存在许多不同的 Stroop 版本，这些差异对临床工作者造成极大的挑战，因为这些版本很可能用的不是同一个程序。举例来说，Salthouse 和 Meinz(1995)发现 3 个不同的 Stroop 测验的差异分数间只有低到中等的相关性。相似的，Shilling(2002)报道在 4 个不同的 Stroop 测验中，其中一个测验干扰的敏感性无法代表其他 3 项测验的干扰敏感性。表现的个体一致性程度取决于采用 Stroop 测验的相似程度。对于临床常用的几种 Stroop 测验，究竟他们之间的分数是否高度相关无从得知。

对一个特定的患者该选用何种版本的 Stroop 测验，有许多因素应列入考量。Stroop(1935)使用了中性和词色不一致的两种情况的差别来衡量干扰程度。若未使用基线(中性)情况，用 Stroop 结果来衡量干扰的解释不值得被采用。不一致状态可测试一般性表现和干扰。若是只给予一种条件，则不可能将这两种特征分开。有时，患者会被给予不一致状态用的卡片，并被要求念出文字忽略颜色，至于第二个条件，他被要求说出颜色并忽略文字。这项比较也

存在问题，因为它混淆任务和干扰间的差异。举例来说，颜色命名与文字阅读的差异既可以反映颜色对文字的干扰，也可反映文字对颜色的干扰。

部分研究者担忧在一致与不一致分开的间隔测验中(比如在 Victoria 和 Golden Stroop 测验中)，可能让受试者有机会发展出一套专注于卡片项目是一个维度还是两个维度的策略。然而，这样的设计形式(不同于 D-K 执行功能系统的混合设计)可能加强了在不同临床情况下(如痴呆)的敏感性，因为这种策略较少被患者应用。相反地，这种不一致测验的间隔呈现可能会使工作记忆的参与最小化，因为这项任务要求在任务中始终保持一贯，使受试者更容易记住自己的目标。每个不一致的词都要求忽略文字，从而强化了目标。

目前有许多方法可以分析测试时间。不同情况下的差异分数常被使用，但不一定能完全独立于一般性的减慢。为了解决基线反应速度的差异，一项将个体基线速度考虑在内的比例分数常被推荐。Golden 假设的模型中，阅读一个 CW 项目的时间是阅读文字所需时间和命名颜色所需时间的总和。相比之下，Chafetz 等研究者(2004)提出阅读一项 CW 的时间反映抑制文字阅读的时间加上命名颜色的时间。这个模型与临床数据较匹配。笔者在使用华山版时发现，卡片 C 的耗时数大于卡片 A 和 B 加在一起的耗时数，即受试者在读字和颜色命名的反应竞争中有时间消耗。

Stroop 的错误分析的研究较耗时数得分的研究少，但在某些版本还是有相关资料的(如 Victoria 版与 D-KEFS 版)。脑部疾病可能损害文字或颜色处理过程或导致分心、速度减慢、重复行为或冷漠。这些不同的特质会影响对错误的感受，因此，在神经系统损害的患者群中分析错误显得格外重要。

华山版的评分指标考虑了错误分析。其指标包括：① 每张卡片的总耗时数、正确数(满分 50 分)。② 干扰量 SIE(stroop interference effects)耗时数＝卡片 C 的耗时数－卡片 B 的耗时数；SIE 正确数＝卡片 C 的正确数－卡片 B 的正确数。③ 语义错误分析Ⅰ型：在卡片 C 中，四种颜色，每种颜色读正确的个数(满分 12 分)与比例，读正确的颜色种类(满分 4 分)。④ 语义错误分析Ⅱ型：在卡片 C 中，四种汉字，每种汉字的干扰程度不同，每个汉字的颜色读正确的个数(满分 12 分)与比例，读正确的汉字-颜色种类(满分 4 分)。⑤ 空间错误分析Ⅰ型：卡片 C 共 5 行，每行 10 个字读正确的个数(满分 10 分)，分析上、中、下空间的差别。⑥ 空间错误分析Ⅱ型：左、中、右空间分别读正确的个数(左 20 个、中 10 个、右 20 个)，分析左、中、右空间的差别。

总之，Stroop 任务的心理学机制包括工作记忆、处理信息的速度、语义激活

和强化反应特征的能力。此试验需要具备与其他执行力测试不同的能力，因此，需要不止一项执行功能测试来进行充分的评估。在不同情况下表现的神经基础并非是确定的，尤其是在不一致的情况下，虽然存在不同的报道说明额叶的重要性，尤其是前扣带回皮质和侧前额叶皮质。虽然额叶系统很重要，但任务的表现依赖于一个更广泛的系统。由于结果可能会受到色觉和视觉敏锐度的影响，因此完好的视力非常重要。Stroop 的测试结果在高龄老年人中应谨慎解释，且在某些年龄组（如>90 岁），这个测试并不合适。最后，Stroop 测验还不能单独用于诊断决策，如果要明确执行功能，Stroop 测验与其他测验指标相结合是必要的。

Stroop 测验有很多版本，其经典版本是受试者阅读色词出现主动反应与自动反应哪种占上方，评估解决两个竞争反应倾向的冲突的能力。Stroop 测验甚至对轻度 AD 患者也具有良好的敏感性。这种敏感性是来源于测验的复杂性和内在难度还是反应选择或反应抑制的特异缺陷，目前还不清楚。反应抑制、竞争反应倾向、习惯化和对新鲜事物的反应构成 MCI 与 AD 认知研究的重要领域。除了经典版本，本节列举了两种其他范式：数字范式，要求读数字的数量而不是数字本身，如表 8-2-2 中是读 5-1-8，而不是读 6-3-1；字母范式，如表 8-2-2 中要求读成 H-L-S，而不是 H-E-H。

表 8-2-2　Stroop 测验版本举例

红	蓝	白
6　6　6 6　6	 3	1　1　1　1 1　1　1　1
H　　　H H　　　H H　　　H HHHHHHHHHH H　　　H H　　　H H　　　H	E E E E E E EEEEEEEE	HHHHHHH H H HHHHHHH 　　　　H 　　　　H HHHHHHH

（郭起浩）

第三节 执行缺陷综合征行为测验(BADS)

执行缺陷综合征行为测验(behavioral assessment of the dysexecutive syndrome, BADS)的目的是为了预测执行缺陷导致的日常问题。整套工具包括测验手册、测试材料、25张评分和评级单。成人版适用于16～87岁的个体;儿童版(BADS-C)适用于8～16岁的青少年。

Wilson等(1996)注意到大多数神经心理测验包括一个明确的任务、一个简短的试验、由测试者激发的任务动机和定义完善的任务完成指标。受试者很少被要求在较长的时间内组织、计划自己的行为或在面对两个或更多的竞争任务中执行主动权,但事实是这样的,执行能力是日常活动的主要组成部分。BADS包括6个测验,它们与真实生活活动相似,会对有执行缺陷的患者造成困难。另外,还有一个执行缺陷问卷(dysexecutive questionnaire, DEX)。DEX可以独立使用,也不是BADS必需的分测验。以下6个测验组成了BADS。

1. 规则转换卡片测验(rule shift cards test) 使用非图片扑克牌来检测受试者对一种规则正确反应的能力和从一种规则转换到另一种规则的能力。因此,这个测试是评估从一种规则转换到另一种规则的能力和工作记忆能力。

2. 程序性动作测验(action program test) 由Klosowaska(1976)最先描述的任务改编而来,要求受试者制定一个五步计划来解决一个新问题。

3. 搜索钥匙测验(key search test) 要求受试者找到他们丢失的东西,然后评估搜索策略的效率。

4. 时间判断测验(temporal judgment test) 由四个关于常见事件的简短问题组成,评估患者进行合理猜测和估计的能力。

5. 动物园分布图测验(zoo map test) 要求受试者说明他们怎样才能到达动物园地图上显示的一系列指定地点,并且遵循一定的原则。用来评估计划能力。

6. 六元素测验修订版(modified six elements) 修改自shallice和burgess 1991年开发的一个任务,要求受试者在有限的时间内组织自己的行动以完成六项任务,不能违反指定规则。

实施指令见BADS手册。一些项目有时间限制。要求测试者在评分单上

记录实施过程的各个方面和评分过程。完整测试大约需要半小时。

一、评分

评分标准见 BADS 手册。每一个测验都有一个评分，0～4 分，整套测验总的得分是每个测验得分相加。因此，每一个完成整套测验的受试者都会得到一个 BADS 得分，为 0～24 分。尽管推荐实施完所有的六项测验来获得 BADS 总分，但也可以在五项测验的基础上按比例换算。DEX 并不是 BADS 正式的一部分，它并不用于计算测验评分，而是可以用来补充测验所获得的信息，提供其他有价值的信息。

总分可以转换成一个标准分，平均数为 100，标准差是 15。这样可以将 BADS 的表现分类为受损、临界、平均水平以下、平均水平、高于平均水平或出众。总分的评价分为 3 个年龄群(40 岁及以下、41～64 岁、65～87 岁)。

二、人口统计学影响因素

年龄影响任务表现，65 岁及以上个体的受试者总分较低。研究者没有报道性别、教育程度、智力的可能影响。对于时间判断测验，有证据表明老年参与者的得分比青、中年人低；性别和语言智力与表现无关。

三、信度

1. 内在信度　Bennett 等(2005)报道认为总分的克隆巴赫系数是 0.60。Gillespie 等(2002)注意到时间判断分测验的项目数太少而不能检测其内在一致性。即使省略了这个分测验，总分的信度也比较低。

2. 再测信度　Wilson 等(1996)报道，29 名正常人在 6～12 个月后再次进行了整套测验(包括 DEX)，以及其他三个执行功能测验(修订的卡片分类测验、认知评估测验和音韵流畅性测验)。再测相关性为－0.08(规则转化测验)～0.71(搜索钥匙测验)，基本低于可接受的水平。总分的再测信度没有报道。很可能是小样本的天花板效应影响了相关性的计算。在其他的执行功能测试中也发现了轻到中度的测试场合的相关性。总的来说，重复测试时的表现有轻度提高的趋势。研究者也报道了检测再测信度的另一种方法。他们计算了两次测验得到相同得分的一致性百分比，一致性百分比也很低。只有两个测验(程序性动作测验和规则转换测验)在两次测试中相同得分达到 70%以上。

在一个患者样本中进行了更短间期的再测，结果相似。Jelicic 等(2001)研究了 DABS 荷兰版(不包括 DEX)的短暂稳定性，使用了一个 22 名成年精神障碍患者组成的样本，所有的患者在间隔 3 周后再次进行 BADS，再测相关性在很低(0.22，程序性动作测验)到很高(0.85，总分)之间。只有一个测验的稳定性系数低于可接受水平(<0.70)。在重复实施中，患者在程序性动作测验中得到更高的得分(大约 1 分)，总分也更高(大约 2 分)。

3. 评定员之间信度　Wilson 等(1996)报道，25 名正常人接受了两个评定员的 BADS 六项测试，评定员之间信度很高(高于 0.88)。

四、效度

1. 与其他神经心理测验的相关性　针对健康个体和神经疾病患者(脑外伤、多发性硬化)，BADS 得分和其他的执行功能测验(如 WCST、TMT - B、迷宫测验、罗伊复制策略、认知估计、言语流畅性)有中度相关性，尤其是总分、程序性动作测验、规则转化卡片测验和所有执行功能测验都有轻到中度的相关性。动物园分布图测验和搜索钥匙测验与另一个检测计划能力的测验(迷宫测验)显示出显著的相关性。时间判断测验与执行测验(包括另一个认知评估)之间的相关性很小，且没有显著性。BADS 测验与记忆测验分数之间几乎没有关系。

2. 临床发现　很多情况下都会出现 BADS 表现受损，它们以不同的方式影响执行功能，包括慢性酒精中毒、精神分裂症、注意力缺陷/多动症。Wilson 等报道，BADS 总分在健康对照者与神经疾病患者(76 名，大多数是近期脑外伤)间差别很大，而且神经疾病患者在 BADS 六项测验中的表现均差于对照者。

BADS 在区分神经疾病(TBI 患者)和非神经疾病患者方面与标准执行功能测验是等同的。在三项测试(程序性动作测验、动物园分布图测验、修订的六元素测验)的总分以及其他的两个测验(迷宫测验、言语流畅性测验)中都观察到有统计学意义的组间差别。BADS 对于区分无脑损伤的个体的正确率很高(84%)，对于神经损伤患者的正确区分度较差(64%)，总体区分率是 74%。其他的执行功能测验也得到了类似的结果。

3. 生态效度　Wilson 等(1996)报道认为执行缺陷问卷评价与六项测验表现之间呈中度负相关性，也就是说，意识到执行缺陷越差，BADS 测验中的执行功能表现也越差，也就是说，BADS 与日常生活中执行问题的客观评分有

关。Bennett 等(2005)评估了有持续性脑损伤的患者,发现各种 BADS 分测验(尤其是程序性动作测验、修订的六元素测验)与临床医生给出的 DEX 评价中度相关。Clark 等(2000)报道,青少年的修订的六元素测验的表现与患者的多动症/注意力不集中显示出中度相关性。Norris 和 Tate(2000)发现程序性动作测验、动物园分布图测验和修订的六元素测验的表现可以预测临床医生对神经疾病患者日常角色功能的评价,而标准化测试(言语流畅性测验、瑞文推理测验)则不能预测角色功能。

(徐　岩　郭起浩)

第四节　威斯康星卡片分类测验(WCST)

威斯康星(Wisconsin)卡片分类测验(Wisconsin card sorting test, WCST)是最常用的执行功能测验,这个测验的用途是评估形成抽象概念、转换和维持分类、应用反馈信息的能力。适用于 5~89 岁的个体。

WCST 是 Berg 和 Grant 于 1948 年开发的,用来评估抽象能力和为了环境应急而转换认知策略的能力。这个测验被认为是评估执行能力的方法,因为它要求策略计划、组织搜索、使用环境反馈信息来转换认知分类的能力、目标指向行为和调节冲动反应的能力。Heaton 等指出这个测验有很多优点,部分是因为它能提供关于问题解决行为的多方面信息,多于任务成功或失败这样的基本指数所提供的信息。基本指数包括持续性错误数、不能维持完整分类数和完成分类数。Heaton 在 1981 年将测试工具和打分流程标准化,并将其作为一种临床工具发表。在更新的手册(1993)中,评分规则得以完善,记录形式被修订,而且提供了年龄在 6 岁 5 个月到 89 岁的常模资料。

测验包括 4 张刺激卡,放在受试者面前,第一张上带有 1 个红色的三角形,第二张上带有 2 个绿色的星形,第三张上带有 3 个黄色的十字形,第四张上带有 4 个蓝色的圆形。然后受试者会拿到两叠卡片,每叠包含 64 张反应卡,反应卡上的图案设计与刺激卡相似,但在颜色、几何图形和图形数量上有所不同。受试者被要求将每一张反应卡与 4 张刺激卡中的一张匹配,每一次选择后都会得到正确还是错误的反馈。分类规则改变时,不会有任何提示。这个测验没有时间限制。标准 WCST 的简化版——WCST - 64,仅仅提供受试者第一叠 64 张卡片。

尽管最常用的是 Heaton 开发的版本，这个测验也有其他的版本，比如改良的卡片分类测验(modified card sorting test, Nelson, 1976)；密尔沃基卡片分类测验(milwaukee card sorting test, Osmon 和 Suchy, 1996)。因为标准版包含一些模棱两可的刺激：可以根据一种以上的类别进行分类，这个缺陷导致测验表现差不是总能被解释。Nelson 在 1976 年移除了可以根据一种以上刺激卡分类的反应卡，从而排除了含糊性。这个测验(WCST－48)包括两叠卡片，各 24 张，仅要求受试者在每一个类型完成 6 次连续的正确分类后即结束。另外，当目标种类变化时，受试者会被告知。许多作者已经提供了这个版本的成人(20～90 岁)常模资料，然而，要注意 Nelson 改良版改变了这个测验的质量。文献已经表明，改良的卡片分类测验和 WSCT 不完全对等，应该被当作独立的测验。尽管关于它与年龄和其他测验分数之间的相关性有一些可靠的证据，但这个版本的再测信度系数很低，表明它对于监测执行功能跨时间改变是不可靠的工具。

密尔沃基卡片分类测验要求患者在分类之前进行表述，从而提供可能对辨别患者形成、保持、转换分类困难有用的额外评分。然而，仍需要进一步的研究来评估这个版本的信度、效度和常模数据。

WCST 可以在计算机上操作完成。Artiola I Fortuny 和 Heaton (1996) 报道计算机版和标准版的结果在正常人群中是非常近似的。同样地，Hellman 等在 1992 年发现计算机版和标准版之间在各种病因所致的精神疾病患者小样本中没有显著差异。但是，两个版本在正常人群之间的得分分布特点尚未发现有均等性。Ozonoff 在 1995 年报道指出自闭症儿童在计算机版中犯错误更少。考虑到文献中指出的两者之间的差异，基于标准版的常模可能并不适用于计算机版。

意大利 Pezzuti(2013)发表了一个新的 WCST 版本(图 8－4－1)，主要改



图 8－4－1　Pezzuti 发表的 WCST 版的卡片举例

变有两点：首先是把形状（三角形、星形、十字形和圆形）维度改为交通工具（汽车、气球、飞机）；其次是测验次数改为 54 次。研究认为，这个新版本不仅接受度更好（妙趣横生），而且对老年人认知障碍的检测也比传统版本更敏感。

一、实施方法

受试者应视觉正常并听到足够的信息，以完全理解测试的结构，且能从视觉上区分刺激卡片的颜色、形状和数量。两叠卡片放在受试者面前。测试员指示受试者把每张反应卡按照他或她所认为应该在的地方堆放在一张刺激卡下面，然后测试员告诉受试者他的选择是正确的还是错误的。受试者会被指导利用这个信息来尽可能多地放对卡片。测试员可以向受试者阐明刺激卡片的意义，以及如何做出反应，但是千万不能通过提供任何分类原则的暗示或从一种类型转换到另一种类型的本质来对 WCST 的整体性造成妨碍。

受试者首先被要求按颜色分类，所有其他的应答都被判为错误；如果对颜色分类达到了 10 个连续正确的应答，然后要求受试者转换分类原则到形状，没有任何警告提示，如果现在还是按颜色分类的话就判为错误。对形式分类达到了 10 个连续正确的应答后，分类原则转换为数量，然后再次回到颜色分类，直到受试者成功完成六次分类（颜色、形状、数量、颜色、形状、数量），或者用完 128 张卡片为止。

记录受试者的表现可能很困难，尤其是当受试者完成很快时。为了简便，记录表共有 128 个反应条目，每一项都是“CFNO”（C 表示颜色，F 表示形状，N 表示数量，O 表示其他），当受试者选择的反应卡与刺激卡匹配时，测试员在对应的一个字母上画斜线来记录患者的应答。为了方便后来的评分，当连续正确的反应达到 10 个，测试员将此次分类连续反应的总数目记录在记录册项目左侧的空白处。如果患者在几个连续正确的应答后做出了一个错误的选择，测试员需从 1 开始重新计数以下的应答。另外，要在 10 个连续正确应答的最后一项下划线，以表明新一轮的正确分类在此线以下开始。

以上检测耗时 15～30 min。

二、评分

WCST 记录册共有 4 页，用来记录患者的信息、患者测试的应答、计算和记录得分。WCST 的评分错误很常见，即使是经验丰富的临床医生。目前国

内使用的基本是采用计算机评分软件来消除记录和评分错误。

评分方法有很多种，见表 8-4-1。比如，完成种类数指在测试中完成的种类数(基于某一分类原则的正确匹配数达到连续 10 个)，计分范围在 0～6，达到 6 后测试停止。当受试者对一个刺激卡的反应发生持续性错误时，以“持续性”来计分。持续性应答数反映了从之前的分类过渡到新分类原则的障碍，或者看见新可能性的障碍。不能维持完整分类数得分是指受试者在完成一个分类过程中，在做出 5 个及以上连续正确的反应后出现 1 个错误选择的次数，它指示着持续使用一个成功策略的障碍。概念化水平百分数指连续正确的反应在 3 个及以上，它反映了对正确分类原则的理解。总应答数、总正确应答数、总错误应答数也要记录。为了有助于研究，持续应答数百分比、非持续性错误数百分比也要计算。但是这些指标并不推荐应用于临床解释，部分原因是这些“百分比”得分的信度低于预期的基本分。

表 8-4-1　WCST 的评分指标与定义

指　　标	定　　义
完成分类数(No. of categories completed)	10 个连续正确配对序列的个数(最大值为 6)
完成第一个分类所需应答数(trials to complete first category)	完成第一个分类的总应答数
持续性应答数(perseverative responses)，持续性错误数(perseverative errors)	受试者对一个刺激类型应答持续错误的项目数
持续性错误百分数(percent perseverative errors)	持续性错误个数除以总测试应答数乘以 100
不能维持完整分类数(failure to maintain set)	患者做出 5 个及以上连续正确的配对，但在成功完成一个类型前做出一个错误应答
概念化水平百分数(percent conceptual level responses)	连续正确应答达到 3 个
学习到学会(learning to learn)	受试者跨连续分类测试中概念化参数变化的平均值；基于相邻两个分类测试错误百分数的差值

在 WCST 得分中有很多都是多余的，比如，完成种类数与总错误数、持续性错误数与持续性应答数是高度相关的，还有一些得分是其他得分的线性组合(如总错误数是持续性错误数与非持续性错误数的和)。在 WCST 中最常用的评估执行功能的数据是完成分类数和持续性错误数。持续性错误数的测量比完成分类数对年龄相关的执行功能下降更加敏感，因此是更好的执行功能衡量标准。

三、人口统计学影响因素

1. 年龄　年龄与其他的人口统计学变量相比，与 WCST 表现的相关性最强，影响测试得分变化值的 20%。从 5 岁到 19 岁，测试表现逐渐提高，在 20 岁到 50 岁之间保持稳定。60 岁以后某方面表现的下降变得很明显。

2. 性别　性别对 WCST 的影响还有争议。大多数作者报道性别与 WCST 的表现无明显相关。Boone 等认为在老年个体中，女性往往比男性表现好。

3. 教育　西方国家的研究表明，受教育水平与 WCST 的得分呈轻度相关性，影响得分变化值的一小部分(4%～7%)，但很重要。在成年人中随着受教育水平的提高，WCST 表现的熟练度逐渐提高。对于儿童，父亲的受教育水平对得分影响很小。

4. 智商　很多作者的文献指出，智商和 WCST 得分有轻度相关性。

5. 种族　母语为西班牙语与母语为英语的个体比较，测试结果相同。

6. 其他　测试表现也受血管健康状况和总体社会经济/健康相关状况影响。

四、信度

1. 再测信度、练习效应和变化识别　对一组 46 名健康儿童和成人分别进行了两次 WCST，中间间隔一个月。调查者报道概化系数(generalization coefficients)在 0.37(持续性错误百分数)到 0.72(非持续性错误百分数)之间。值得注意的是，在正常儿童中 WCST 标准得分的标准误很大。因此，对正常儿童进行再测时很可能会得到不同的结果，对 WCST 得分的解释应该非常谨慎。Paolo 等的发现与此相似。他们在大约 1 年时间的间隔后再测试了 87 名正常老年人，每名参与者在两次测试中的认知功能正常(DRS 大于 130 分)。结果发现稳定系数很低，总错误数范围为 0.12～0.66。大多数受试者在再测时表现更好，再测时 5 项得分平均提高 5～7 个标准得分点。Basso 等在间隔 12 个月后再测 50 名健康人群(年龄均值为 32.5 岁，标准差为 9.27)，发现再测结果的所有指标几乎都显著改善，尤其是完成测验所需应答数。完成第一次测验平均需要 101.12 张卡，12 个月后仅仅需要 84.74 次应答，持续性错误和应答数下降一半。练习效应的存在和效应大小在平均及平均水平以上智商的个体中是相似的。

有研究者认为一旦记忆完好的个体掌握了测试的分类方法和转换原则，WCST就不再能衡量问题解决能力了。这表明对于正常人来说，WCST的稳定性低，再测时，不再能以相同的方式衡量问题解决能力。并不是因为规则改变的具体过程被记住了，而可能是测试需要的知识、有效的测试策略被获得了，从而提高了后来的表现。然而，在临床样本中，信度会高很多，提示该测验对于临床人群更加敏感。WCST－64的稳定性高于标准版WCST，接受标准版测试的个体在一次测试后会有两倍的经验。需要牢记在心的是，有意义的参数变化给临床医生提供了执行功能改变已发生的证据，但是WCST固有的某种程度的不可靠性，意味着即使参数变化在阈值以下，我们也不能确定没有任何变化发生。

2. 评定员之间信度　计算机软件的使用减少了评定员的人为误差，提高了信度。

五、效度

1. 测试内相关性　大量多人群（如混合精神病或神经病患者、精神分裂症患者、创伤性脑损伤患者、脑血管意外患者、正常老年人）的因素分析研究已经发表，大多数使用的是探索性因素分析，已经发现了支持两种以上因素的证据，其中3种因素已经确定，这些认知过程包括转换分类能力、问题解决/假设检验和应答维持。

最近的一项大型验证性因素分析，使用的样本包括620名神经疾病患者、228名精神疾病患者和373名非临床对照，它支持探索性因素分析文献中报道的3个因素。然而，只有反映整体执行功能的第一个因素（包括持续性应答数、概念化水平百分数、完成分类数、总正确数）有统计学合理性。第二类因素（a：概念化水平百分数、完成分类数、总正确数、非持续性错误数；b：总正确数和不能维持分类数）很可能反映的是有意义的认知能力，稳定性比较低，只有当所有的受试者都用完128张卡片才有稳定性。也就是说，相比总体执行能力，WCST对于认知过程更敏感，尽管它没有完美地评估这些过程。Greve（2005）建议使用两种结束标准（即完成6类测试或用完128张卡片）可能会有利于不同的因素分析方法。也有建议所有的受试者都用完128张卡片，因为过早停止会导致一些有用信息的丢失。

2. 与其他神经心理测验之间的相关性　在正常儿童中，WCST的持续性与韦氏儿童智力量表修订本和全量表智商的相关性为0.30～0.37。相比之

下，在异质性精神疾病样本中，全量表智商、行为智商而不是言语智商，与WCST的表现轻度相关。在健康成人，总错误数与流体智力关系密切。WCST评估的是一些与心理测量智力有关的总体推理能力。它与记忆、注意力/工作记忆评估之间的关系有争议。比如，一些研究者报道持续性指数与注意力/工作记忆和情景记忆之间轻度相关。Paolo等分析了WCST得分和注意、记忆评估的关系，结果发现在正常老年人和PD患者中，WCST得分不随记忆和注意力/工作记忆得分而变化。

Miyake等将基本执行功能分为定势转移、信息更新、优势抑制。WCST表现可由“转移”能力有效预测。一旦将“转移”能力考虑在内，“抑制”和“更新”能力将不能对WCST表现做出任何解释。这些发现说明我们反复提及的执行功能，即转移能力，大大影响WCST表现，这为WCST是一个执行功能测验提供了支持。

3. *与其他执行功能测验的关系* WCST评估的是概念化处理的范畴。基于多种临床样本下对WAIS－R分测验的因素分析，Golden等认为分类测验反映了空间分析技巧，而WCST评估的是不同于其他测验的要素。Perrine注意到WCST与属性识别有关，包括对相关特征（如正确的答案基于颜色、形状和数量的刺激属性）的辨别。在分类测验中，正确的答案基于识别正确的原则或规则，而不论刺激的属性如何。Perrine注意到这两个过程中有30％的区别，虽然说两个过程有很多重叠，但他们反映的是概念化推理中的不同能力。Minshew等观察到WCST不完全是一个概念识别任务，因为测试对象并不知道相关规则发生变化，这就对自发评价和假设形成提出了要求，两者可被持续性错误得分评估。Macpherso等发现，衰老会影响WCST和其他依赖于额前区背外侧功能的测验，但是对于依赖于腹内侧区的测验几乎没有影响。

六、临床研究

本书侧重归纳WCST在中老年人群中的运用，而儿童人群中的应用资料不在此列举。Milner(1963)在她对WCST的经典研究中发现，额叶背外侧切除的患者和眶前额叶及后部病变的患者有明显不同。额叶背外侧病变的患者表现为“从一种分类原则转换至另一种，明显是由于之前反应模式的保留性干扰”的能力缺陷。一些后来的研究使用了功能影像学（PET或SPECT）和MRI成像，也支持WCST对额叶功能敏感的看法。最近一篇比较额叶损伤和脑后部损伤患者的meta分析表明，额叶损伤测试的表现更差，尤其是背外侧

损伤或急性外伤(1年之内)的个体。然而,Demakis注意到虽然效应值很大(背外侧损伤的患者,$D=-1.3$),分数重叠也很多(大约35%)。这也就是说,在准确区分额叶和非额叶病变人群时没有足够的间隔。其他人也注意到一些广泛额叶病变的个体测试表现正常。另外,meta分析发现局灶皮质切除的患者音位流畅性(phonemic fluency,如列举F开头的单词)比起WCST得分,与额叶病变存在的关系更密切、更有特异性,这与1296名创伤性脑损伤患者的30个研究的meta分析结果相似,后者表明音位流畅性对于创伤性脑损伤患者来说比WCST更敏感,可能对于这类人群是更好的执行功能障碍评估。另外,meta分析显示右侧和左侧额叶损伤的受试者,其WCST得分比较没有明显差异。

创伤性脑损伤、孤独症、多发性硬化、帕金森病、强迫症、科萨科夫综合征、注意力缺陷/多动症/冲动型、大量饮酒、产前酒精暴露和长期使用可卡因及多种药物均导致WCST表现不佳。

衰老相关表现下降在WCST中非常明显。一些人提出测试表现的年龄差别与年龄相关的工作记忆衰退有关。Fristoe等对青年人和老年人进行了WCST、工作记忆和处理速度评估,结果表明WCST表现的年龄差异与工作记忆和反馈信息的有效处理都有关系。他们提出,年龄相关的处理速度下降导致WCST表现不佳是由于限时和限同时性机制,即处理速度下降导致连续执行实施更少,先前的处理结果可能并不适用于之后的处理。Hazlett(2015)比较了AD的认知正常的一级亲属与没有AD家族史的对照组的WCST表现,发现分类数、概化水平百分比、持续性错误数与非持续性错误数均有显著差异,在控制年龄、性别、教育等元素后,仍然是有家族史的一组表现得更差。

早期有研究认为精神分裂症患者显示出WCST表现受损,但是当考虑到语言智商时,患者与对照者的差异不明显,表明WCST缺损在精神分裂症患者中反映的是更普遍的智力缺陷,而不是执行功能的障碍。但是,最近的研究认为精神分裂症患者存在执行功能障碍。Vogel(2013)发现精神分裂症患者进行WCST检查时,开始4张卡片的准确性可以预测完成全套测验的分类数与持续性错误百分比,他认为患者阴性症状与难以恰当使用负反馈相关(这也说明阴性症状不继发于焦虑症、抑郁症)。

另外,有文献表明抑郁症或焦虑症会影响WCST的表现,尤其反映在持续性,不能维持完整分类数增加,概念化水平百分数下降。抑郁症与问题解决能力下降有关,症状越重表现越差。这暗示神经心理测验应该包括对情绪紊乱的评估。如果抑郁或焦虑症状出现,临床医生不能总结说WCST表现异常

就预示着某种神经功能障碍，因为在情绪障碍的情况下可能会表现差，即使没有任何明显的神经疾病。

需要注意的是，精神疲劳也会导致执行能力下降。工作时间长的健康青年比起非疲劳个体，完成种类数更少，持续性更高。Ozonoff 指出在 WCST 中表现良好也要求一定水平的社会意识和动机来参与语言反馈。这是评估社会意识和动机缺陷患者时应当考虑的重要方面。

帕金森病所致痴呆(PDD)的前驱期存在轻微的执行功能障碍。Woods 等在随访患者一年之后，发现 WCST 持续性在识别后来发展为痴呆的 PD 患者身上显示出一些诊断分类精度(总预测能力为 68%)。相似地，对有严重情绪障碍风险人群的长期随访表明，WCST 表现受损与后来双相障碍的发展有关。

七、简化版

虽然使用简化版会降低信度，但 WCST－64 与完整版对应的得分高度相关。然而，在成年人具体测试时，WCST－64 的结果与完整版并不都一致，但对于区分表现受损(<40 T 分)还是不受损的一致性很高(86%分类相同)。Merrick 等(2003)发现在一个持续创伤性脑损伤成年患者样本中，有 1/4 WCST－64 持续反应 T 分低于相应的 WCST 变异版本 10 分以上。这个发现随年龄而变化，比起青年人，老年人的 WCST、WCST－64 测试结果显示出更强的一致性。

有一些证据表明简化版对 AD 和 PD 的功能缺陷敏感。在创伤性脑损伤后的急性恢复期(中位时间为损伤后 34 天)，完整版和简化版对于损伤严重程度的敏感性相同。WCST 对于创伤性脑损伤后的后急性期(中位时间为损伤后 3 个月)损害比 WCST－64 更加敏感，因为更多比例的此类个体在完整版 WCST 中的得分受到影响(58%vs. 32%)。这表明对于疾病更轻微的患者，完整版更合适。然而，WCST－64 与 WCST 在预测创伤性脑损伤患者出院后的功能障碍水平方面是一致的。

八、评价

WCST 是评估认知功能下降最常用的方法。记住这一点很重要——执行功能是可分割的，不同的执行功能成分有助于不同的复杂执行任务的完成。因此，仅仅依赖一个任务，如 WCST 作为执行功能的总体评估是不充分的。

成功完成 WCST 背后的认知过程是复杂的。这个任务要求很多技能，包

括基本的视觉处理能力、数字能力、规则归纳能力、识别最相关刺激属性的能力、快速处理能力、维持目前分类的能力、心理定势转换的能力和适当的动机集合。定势转移能力可能是 WCST 表现优劣的核心。然而，除了定势转移能力，其他处理过程的缺陷也会使 WCST 表现受损。也就是说，在这个复杂的测试中，患者有很多种可能方式会显示出表现受损。

同样，在这个测试中表现受损可由于多种病因，并不都与额叶功能相关。WCST 对额叶功能敏感，但并不能用于预测局灶性的额叶病变。因此将 WCST 看作是个要求多处神经元网络功能的多因素测试，貌似更可信。另外，有证据显示，其他的测验（如音位流畅性）可能提供对执行功能更好的评估，因为它们对额叶损伤更加敏感和有特异性。

使用者应该记住，WCST 变异版本的信度在正常人群中很低。信度在临床样本中要高一些（如持续性错误），但不是所有的 WCST 指标都如此。因此，临床医师可能会被错误信息误导，所以在分析个体得分或测试结果的特征性廓图时，应该非常谨慎。

有很多 WCST 的结果是多余的，建议临床医师使用 WCST 有代表性的变量（如持续应答数、不能维持完整分类）。完成种类数和持续错误数都可在年龄相关认知变化研究中被当作预测指标或执行功能的评估值。Rhodes 认为持续性错误比起完成种类数对于年龄差异更加敏感，如果只使用 WCST 中的一个得分，前者可能是认知功能更好的指标。

对于一些患者完成 WCST 可能要花很长时间，因此，他们对简化版使用的兴趣越来越大，特别是 WCST－64。然而，考虑到一些研究中两者之间的分数缺乏一致性，所以使用简化版时要小心。简化版使用受限的其他原因还有它们不能解释测试过程中出现的表现形式。WCST－64 可能仅在年龄分布的两极（如≤16 岁的儿童和≥60 岁的成人）是完整版的合理替代者。然而，青年人在完成第二叠卡片时会“领悟”。有证据表明对于轻微功能障碍的患者，完整版可能更适合。同样地，当实施完整版时，有两个结束标准：完成 6 个分类或用完 128 张卡片，看哪个在前。Greve 建议常规使用用完 128 张卡片，因为过早结束会使有用的信息丢失。

计算机版也是可用的。然而，当基于计算机版的得分得出结论时临床医生和研究者应当谨慎。另外，测试老年人时，计算机版和纸质版的差异更应考虑，因为老年人的计算机技巧更差、更不适应。

（徐　岩　郭起浩）

第五节 范畴测验(CT)

范畴测验(category test, CT)的目的是评估受试者的抽象或概念形成能力,面对复杂和新问题解决时的灵活性,以及经验学习能力。

成人版(booklet category test, BCT)包括一本专业手册(professional manual)、两卷刺激板(two volume set of stimulus plates)和50本反应手册(50 response booklets)。

电子版允许测验者实施3个版本中的任意一个:标准208项版和两个简略版,其中的一个120项,另一个108项。

范畴测验由Halstead于1947年开发,用以评估概念化特性的能力,如尺寸、形状、数量、位置和颜色,它也是整套Reitan测验(Halstead-Reitan范畴测验,HRCT;Reitan和Davison,1974)的一部分。最初的版本包含336项,编制为9个分测验。1948年,Reitan将其减少为7个分测验,总共包含208个项目,并通过幻灯片投影的方式呈现出来。每一套都基于不同的原则,如物品的数量、奇数(odd)刺激物的空间位置等。例如,第一套显示罗马数字Ⅰ~Ⅳ,患者被要求在反应键上指出测验设计所要求的阿拉伯数字(1~4)。受试者对每个分测验的项目进行猜测,并需要从他们收到的正确或错误的反馈中来猜测分测验的原则,没有关于测验原则的提示。因此,测验需要通过反应-可能(response-contingent)反馈来推测分类原则,并在它有效的时候使用该原则,在它不再有效的时候舍弃该原则。成人版适用于大于15岁6个月的个体。中间版适用于9~15岁6个月的个体,包含168个项目,分为6个分测验。儿童版包含80个项目,分为5个分测验,适用于5~8岁的儿童。

幻灯片版测验的一个问题在于它依赖昂贵的、笨重的设备,很难应用于床旁测验。测验的手册版和计算机版更便捷。有一些证据表明,这些形式得到的结果与最初的幻灯片版是等同的。手册版很流行,最近很多文献都使用成人手册版。Nici(2013)针对计算机版与原版进行比较研究,发现差异在2分以内,计算机版可以代替幻灯片版。同期,Silk-Eglit(2013)也开发了一个计算机版的范畴测验,与传统版相比,所有的分析指标(总分、错误分、分测验错误分)没有显著差异。

另一个问题是认知受损的受试者完成测验需要很长时间(长达2 h),致使

范畴测验的成人简略版应运而生。简略版范畴测验的手册形式包含100个项目，有一些证据表明简略版范畴测验在心理评估方面的特征识别能力，以及与标准神经心理测验的关系都和HRCT相似。它的得分也和手册版相似。

一、实施方法

1. *手册版方法*　向受试者呈现测试项目的仪器包括一个幻灯片投影仪、一个有显示屏的控制台和测试者的控制板。受试者必须发现潜在的原则，从而从4个刺激物中选出1个。受试者通过按4个编号为1、2、3、4的刺激物中的1个做出反应，4个刺激物从左向右排列，放在屏幕下方。

测试项目被投影出来，首先呈现的是8个罗马数字项目，从Ⅰ～Ⅳ。因为受试者反应板上的按键是数字，第一个分测验就用来在条目和受试者反应键之间建立联系，以及让受试者熟悉测验过程，减少测试焦虑。正确的反应会得到一个钟声，错误的反应会得到一个蜂鸣器声。在每个分测验之间会有一个空白框架。

在测试者控制板的左边有一个总开关，打开开关，屏幕和测试者的控制板就可以使用了。控制条目呈现的手控开关在测试者控制板的右边。当快速按下这个开关，新的条目就会取代屏幕上之前的条目，不要一直按这个开关。对应受试者4个反应键的四路开关在测试者控制板的中间。为了给受试者反馈，测试者必须在每个条目中都将开关拨到合适的数字位置，在记录单表上可以查询。测试者必须在更换幻灯片之前将答案键拨好。

作为一条基本原则，当受试者认为必要时，指令的任何一部分都是可以重复的。这样做的目的是使受试者清楚地理解他(她)所面对的问题以及解决的规则。不能告诉受试者具体原则，但是当遇到概念形成能力严重受损的受试者时，要督促受试者仔细研究图片，要求他们描述刺激物，然后测试者可以这样提问:“你知道正确答案是什么了吗?”督促他们试着注意并且记住图片是如何变化的，因为这常常会提供测试原则的线索；试着去思考回答正确时的原因。然而，测试者不应该提供关于测试中问题解决方法的信息。唯一的信息来自每个反应之后的钟声或蜂鸣器声。

认知障碍患者有时会觉得这个测验很费劲而且令人沮丧，测验者必须尽所有努力来鼓励受试者继续测验。如果受试者在任意一个分测验(Ⅲ～Ⅵ)的前20个项目中没有任何进展，这时最好打断受试者，并按比例分配错误得分，这样比冒着不能完成测验的风险更合适。

另外还有一些要注意的地方：① 虽然速度不是一个影响因素，受试者也不应该被催促，但是他们也不被允许在过长的时间内做出反应，因为会影响测验的连续性。② 测验者应该一直注意屏幕上的幻灯片，不应仅仅关注受试者的表现，还应该确保幻灯片播放顺序的正确。③ 测验房间的光线应该适当暗一些，只要足以使测验者记录错误即可。④ 受试者应该站在屏幕的正前方，因为从侧方很难辨别颜色。

2. 实施时间　成人全套，包括常用的 BCT，需要大约 40 min，但是认知损害的个体可能长达 2 h。简略版（SCT），包括儿童版（CCT），需要 20 min。

3. 评分方法　记录总错误数。错误粗分可以转换为量表分，然后转换为人口折算 T 分。

二、人口统计学影响因素

1. 年龄　随着儿童发育成熟错误数下降，尤其是 10 岁以后，但随着年龄的增加错误数增加。Heaton 等分析，年龄是引起测验得分最大的变异，其次是教育。

2. 智商/教育　教育水平越低，得分越低。但是也有报道认为教育并不影响测验得分。

3. 性别　性别对测试得分几乎没有影响。

4. 种族/文化　表现受文化水平的影响，文化水平低的个体得分低。

三、信度

1. 内部一致性　使用奇偶对半信度法和 α 系数来计算范畴测验 208 个项目的内部一致性。正常成年人和脑损伤样本总分的信度系数都很高（>0.95）。分测验Ⅰ（$r=0.46$）和Ⅱ（$r=0.65$）的信度系数不佳，其他的分测验信度系数很高（0.77～0.95）。Booklet 版、SCT、CCT 的总分半信度系数稍微低一些，但是可以接受（SCT 为 0.81，CCT 水平 1 为 0.88、CCT 水平 2 为 0.86）。分测验Ⅰ、Ⅱ到Ⅲ的难度水平明显提高，事实上，有证据表明分测验 I、II 的项目太简单，有天花板效应。

2. 再测信度　在正常个体中，短期间隔（3 周）和长期间隔（1 年）的再测信度呈中等到良好。对于严重神经损害患者，范畴测验的再测信度很高，信度系数在 0.90 以上，即使相隔两年进行复测。儿童和精神分裂症患者的相关系数有些低，为 0.63～0.75。长期随访一组学习障碍儿童的结果显示为轻度相

关性(相关系数大约 0.4),同时错误得分下降,解释这个结果应该注意到第二次测试从中间版变成了成人版。

3. 练习效应 相关系数仅仅表明受试者在两次测验中的排名是否在相同的名次,而不是受试者得到了相同的分数。事实上,有意义的变化,或者说练习效应,即使是在中度认知损害的患者中也会出现。因此,第二次测验的分数没有提高可能是异常的提示。年轻、教育程度高、初次完成者、测试间隔短者的分数提高得更多。另外,表现倒退到平均水平者也可出现,初次得分差者第二次测验的分数可有一个很大的提高,而初次得分高者第二次测验的分数提高很少甚至倒退。

正常成年人在短期间隔(3 周)之后进行再测试,平均最初得分的变化值为 23.5 分($SD=18.5$),相对于初次表现的平均变化率约为 46%。Dikmen 等(1999)报道发现在一个 384 名正常人和病情稳定的个体组成的样本中,9 个月后再次测验时的错误得分平均下降约 10 分。

四、效度

1. 与其他测验之间的关系 范畴测验的总分(成人版和儿童版)与全套量表的智商评定中度相关,尤其与韦氏智力量表的操作分测验得分相关。范畴测验并不能区分非语言推理之外的能力。

范畴测验与学习记忆神经心理评估轻度相关,这为范畴测验是一个学习能力评估指标提供了一些证据。然而,范畴测验并不将学习当作一种单纯的能力概念来评估。范畴测验与学习成绩轻度相关(r 为 0.14~0.38)。

2. 因子分析 虽然范畴测验是由几个分测验组成的,但临床分析和大多数研究倾向于用一个总分来表示。普遍认同的是,范畴测验是一个复杂的测验,评估很多不同的因素。使用很多不同样本(比如精神分裂症患者,脑损伤患者,对照组)的研究表明范畴测验至少评估两个因素:空间位置推理(分测验Ⅲ、Ⅳ、Ⅶ)和比例推理(分测验Ⅴ、Ⅵ、Ⅶ)。评估空间推理因素的分测验受年龄的影响,而评估比例推理因素的分测验更多地受教育的影响。该测验对脑外伤的严重程度很敏感。

五、临床研究

范畴测验对于很多大脑疾病都很敏感,在评估神经损伤方面是否存在敏感性几乎与全套 Halstead-Reitan 相同。然而,其他依赖于记忆和处理速度的

测验(如听觉词语学习测验)对于各种脑损伤和(或)脑外伤严重程度方面的敏感性更高。将范畴测验分测验的得分汇总成一个总分可能会限制它的敏感性。

范畴测验表现不佳并没有显示出与脑损伤的特别定位或侧重有一致的关系,尽管它最初是被设计用来检测额叶损伤的。抑郁症和精神分裂症患者的表现也会受损。

在预测效度方面,Barreca 等(1999)发现范畴测验过程中犯错误很少的卒中存活者,其肢体功能恢复最好。

六、评价

范畴测验在神经心理学中的应用时间很长,目前仍然是一个很流行的测验。Lees-Haley 等(1996)发现范畴测验在进行司法评估的神经心理学家所实施的测验中占到 32%,在常用的测验中位列第八。Camara 等(2000)发现范畴测验在神经心理学家最常使用的评估工具中排在第九位。Rabin 等(2005)最近报道范畴测验在最常使用的执行功能评估方法中排列第三。由 De Filippis 提出的手册版比原始幻灯片版使用更广,因为它相对便宜而且便捷。

有证据表明范畴测验在不同的媒体版本(幻灯片、手册、电脑)中具有稳健性。计算机版的优点包括测验实施无误和附加的数据收集,如反应时间和持续计数。然而,随着计算机相关焦虑的个体增加,错误得分和反应延迟也会增加。Heaton 等(1991)提出的常模更加适用于计算机版,然而将其应用于有计算机相关焦虑的个体就会有问题。也许随着计算机的普及,计算机相关焦虑的个体会减少。

虽然 Reitan 和 Wolfson 认为将初始分数根据人口变量进行转换可能并不是一个合理的处理,但是现有的文献表明考虑人口统计学背景的常模能更准确地反映患者的神经心理状况。相对的优点主要是将正常个体错判为认知功能受损个体的可能性减少。

范畴测验是一个多因子测验,将分测验得分汇总为一个总分会限制诊断信息。成人版和儿童版都已被发现范畴测验可以分为两个因子,但是常模数据只有总分,并不适用于这些基于两个因子的得分。

当再测验时,范畴测验有很大的学习效应。如果再测验时没有进步说明存在问题和(或)没有尽力。如果有大量错误也是可疑的,特别是在分测验 I

和Ⅱ中。

七、范畴测验和 WCST 的比较

范畴测验和威斯康星卡片分类测验(Wisconsin card sorting test，WCST)在某种程度上都需要通过反应回馈来对分类原则进行推理，当一个原则有效的时候继续使用，在它不再有效时放弃。然而，这两个测验并不相同，仅显示出轻度的共同变异(common variance)。Perrine(1993)认为两个测验与概念形成的不同有关。WCST 与属性辨别有关，包括相关特点的区分，而范畴测验与原则学习更加相关，它评估分类原则的推理能力。类似地，基于对 WAIS-R 分测验的因素分析，Golden 等(1998)认为范畴测验反映的是空间分析能力(依赖于操作分测验)，而 WCST 评估的因素并不依赖于其他测验。范畴测验所需的感知抽象能力比 WCST 所要求的更难。另一方面，WCST 需要受试者意识到正确的匹配原则会在没有警告的情况下定期变化，范畴测验没有类似的需要。WCST 提供对持续性趋势的评估，而范畴测验没有。Adam 等(1995)报道在乙醇(酒精)中毒的患者中，HCT 的分测验Ⅶ与扣带回和额叶背外侧、眶回内侧区的葡萄糖代谢程度有关。WCST 的完成种类数仅与扣带回的葡萄糖代谢有关。另外，Anderson 等(1995)分析了 68 名创伤性脑外伤患者的 MRI 扫描，发现范畴测验和 WCST 的表现都与脑外伤有关，但与局部额叶损伤的体积、有无额叶损伤和非特异性结构性变化(萎缩)的程度均无关。

使用两个测验中的哪一个取决于要诊断的问题。比如，如果评估者想检测持续性趋势，就应该选择 WCST。如果检测者想要一个更难、更敏感的抽象能力测验，那么范畴测验是更好的选择。因为有实施顺序的影响，临床医师应该考虑哪一个工具能提供关于问题最好的信息，那么就应该使用适合的那个测验。

(徐　岩　郭起浩)

第六节　加利福尼亚卡片分类测验(CST)

Delis 等于 1992 年根据经典的威斯康星卡片分类测验(WCST)的设计原理编制的修订版，命名为加利福尼亚卡片分类测验(California card sorting test)，临床用于执行能力的测验，主要反映优势抑制能力，即将注意力集中在

相关信息及处理过程上，抑制无关的信息和不适宜的优势反应，其具有代表性的总评分指标已经在神经精神疾病中广泛应用。

该测验共有 3 套卡片组成，每套为 6 张卡片，每张卡片上均有一个单词，可根据 6 种不同的分类方法将一套卡片平均分成两组，其中 3 种方法是根据卡片上面的字分类，另外 3 种分类方法与卡片上的字无关(表 8-6-1)。标准的测试过程包括以下 3 个方面：

(1) 检查受试者的自发分类。将一套的 6 张卡片放在受试者面前，随机排列，嘱受试者将其分为两组，每组 3 张，并告知有不同的分类方法。当受试者完成一次分类后，询问其分类依据。之后，让受试者换一种分类方法将这 6 张卡片分成两类。接着给受试者 3 分钟时间，嘱其快速用尽可能多的方法将这 6 张卡片分成两类。

(2) 检测受试者的描述分类。由检查者将卡片正确分成两类，嘱受试者指出分类依据。

(3) 检测受试者的线索分类。受试者需根据检查者的分类提示将 6 张卡片分成两堆。检查者首先给出相对抽象的线索(如“请根据颜色分类”)，若受试者无法完成，检查者再给出更详细的线索(如“这些卡片是绿色的还是黄色的”)。

表 8-6-1　加利福尼亚卡片分类测验目标分类方法

分　类	第一套	第二套	第三套
文字分类	地上的物体；空中的物体	陆栖动物；水栖动物	宗教；服装
	运输工具；生物	家养动物；野生动物	单词以“F”开头；单词以“B”开头
	复合词；非复合词	大动物；小动物	抽象的词；具体的词
非文字分类	大卡片；小卡片	黑三角形；白三角形	大写字母；小写字母
	绿卡片；黄卡片	三角形在字上；三角形在字下	圆形；三角形
	直线边缘；曲线边缘	背景线左斜；背景线右斜	单词下划虚线；单词下划实线

一、评分

原版的加利福尼亚卡片分类测验分别从自发分类、描述分类和线索分类三个方面进行评分。

1. 自发分类　共有 9 项评分：① 总分类数，即受试者在 3 min 内完成的

分类数；② 正确分类数，即符合表 8－6－1 中分类法的数量，每次计 1 分(最高 18 分)，包括文字分类得分和非文字分类得分(最大分值各为 9 分)；③ 非目标分类数，即不符合表 8－6－1 中分类法的其他分类法数量；④ 非平均分类数，即受试者未将 6 张卡片平均分成两类的次数；⑤ 重复分类数，当受试者重复做某一种分类时，该重复分类不计分；⑥ 正确描述数，若受试者能正确指出两部分的特点则计 2 分(如“这些是黄色的，那些是绿色的”)，若受试者仅描述某一部分则计 1 分(如“这些是黄色的，那些不是”)，最高得分 36 分，其中文字分类部分 18 分，非文字分类部分 18 分；⑦ 非匹配分类数，即不完全符合表 8－6－1 中的分类法但又相关的分类数；⑧ 错误分类数，即与表 8－6－1 中的分类法完全不相关的分类数；⑨ 持续分类数，即在本次分类时说出前一次分类方法的次数。

2. 描述分类　评分与自发分类的⑥～⑨项相同。

3. 线索分类　共有 5 项评分：① 准确分类数，即受试者按照检查者的提示正确完成的分类数，根据抽象线索完成时计 2 分，根据详细线索完成时计 1 分，总分最高 36 分(文字部分和非文字部分各 18 分)；② 错误分类数，即与表 8－6－1 中的分类法完全不相关的分类数；③ 非平均分类数，即受试者未将 6 张卡片平均分成两类的次数；④ 持续分类数，即在本次分类时说出前一次分类方法的次数；⑤ 重复分类次数。

Fossati 等在评价单相抑郁和精神分裂症年轻患者的执行功能时，所采用的每套含 8 张卡片的 8 种分类方法(5 种空间分类和 3 种文字分类)，测验自发分类的限定时间为 5 min，评分时选取了原版的大部分指标，包括自发分类的总分类数、正确分类数(最大 24 分)、正确分类的概率、正确描述的概率、重复分类数、重复描述数；描述分类的正确描述数、重复描述数；线索分类的抽象线索时正确分类的概率、具体线索时正确分类的概率和重复分类数。

在临床应用中，研究者往往简化了该测验的分析指标。Greve 等在分析儿童和青少年执行能力时采用了以下指标：① 自发分类中的总分类得分；② 自发分类中的正确分类得分；③ 自发分类中的正确描述得分；④ 描述分类中的正确描述得分；⑤ 线索分类中的正确分类得分。

根据加利福尼亚卡片分类测验翻译修订的中文版在上海市区初中及以上文化程度的健康老年人群中的应用结果显示，平均总评分为 3 分(满分为 8 分)，相当一部分受试者呈现“地板效应”，说明该版本对于中国老年人群来说难度过高。为了编制适宜中国老年人群的简短而有效的判断执行能力的测验

卡片，笔者开发了中文卡片分类测验（CCST）。该测验分别由图形、词语和数字共3套卡片组成，每套为6张卡片，要求按照形状、数量、颜色和意义4个维度进行自发分类。操作步骤如下：每次提供6张卡片，随机排列，嘱受试者将其分为两组，每组3张卡片，受试者需说出这样分类的理由；然后，记录所需要的时间，检测者不能告诉受试者共有多少种分类方法，也不能暗示分类原则。如果受试者仅完成1～3种分类，可以说“还有吗?”，给予鼓励，但不能告知分类方法；只要求自发分类，不要求结构分类和线索分类。若在受试者所回答的答案中无标准分类方法，视合理性计正确或错误评分，但每套卡片总评分≤4分，3套卡片的总评分为12分。CCST的临床适用性良好，几乎无“地板效应”和“天花板效应”，受试者回答无歧义、完成率高、耗时少（平均约为5min）；其检测效度亦较高，当CCST总评分≤6分时，鉴别阿尔茨海默病的敏感度为94.32%，特异度为74.24%；当CCST总评分≤5分时，鉴别阿尔茨海默病的敏感度为74.35%，特异度为89.52%。因此建议，当CCST用于受教育程度较低（小学和初中文化）的受试者时，以选择“CCST总评分≤5分”作为临界分值为宜；而用于受试教育程度较高（高中和大学文化）的人群时，以选择“CCST总评分≤6分”为临界分值更佳。

二、评价

目前国内外应用加利福尼亚卡片分类测验分析执行功能的不同方面，如概念的形成、提取以及解决问题的能力。与威斯康星及其他卡片分类测验相比，加利福尼亚卡片分类测验有其独特的优势。第一，将概念的形成分为文字的和非文字的，可分别从两个角度进行评价。第二，可测试患者解决问题能力的极限，进一步对认知能力做出解释；Beatty和Monson认为该测验用于帕金森病患者执行能力时较WCST更敏感。第三，当患者在该测验中回答错误时，检查者并不给出提醒，因此患者的体验相对较好，不容易产生厌恶的情绪。Crouch等研究认为高度偏执的右颞叶切除患者往往不能顺利完成WCST，而对该测验的耐受性较好。

（王　萍　郭起浩）

第九章

社会认知的评估

社会认知(social cognitive)是个人对他人的心理状态、行为动机、意向等做出推测与判断的过程。社会认知的核心组成部分是心理理论(theory of mind, ToM),它是一种推测他人心理状态或根据他人心理状态推测其行为的能力,这种能力在人际交往、学习及社会适应等方面有着重要作用,该理论最初由 Premack 和 Woodruff 提出。

美国精神疾病诊断与分类手册第五版(DSM-5)把认知领域区分为6个,除了记忆、语言、注意、执行、空间外,还有社会认知功能,后者包括:情绪识别,辨别面部图像中的情绪,这些面部图像表现各种积极或消极的情绪;心理推测,能够考虑他人心境或经历的能力;行为调控,对于失抑制和冲动的评估。

Sachdev 等发表的血管性认知障碍(VCD)的认知评估领域也包括社会认知(情绪识别、社会抑制、心理理论、移情)。常用的社会认知包括眼区阅读测验(the reading the mind in the eyes test, RMET)、失言测验(faux pas text, FPT)、一级错误信念任务测验(the first-order false belief test)、二级错误信念任务测验(the second-order false belief test)、爱荷华博弈任务(Iowa gambling task)、故事完成测验(story completion task, ciaramidaro)、酒店任务测验(the hotel task)、Yoni 测验(Yoni test)等。本章介绍眼区阅读测验(RMET)、失言认知测验(FPT)和爱荷华博弈任务(IGT)。

第一节 眼区阅读测验(RMET)

RMET 常用于研究社会和情感技能,是最常用的心理理论测验之一,它简

单而实用，无暗示风险，被广泛应用于不同文化背景及不同的人群中。

RMET是一种通过面部表情，特别是眼睛区域以评估受试者理解别人感情、思想、意图的能力的测试，早在1997年由Baron-Cohen等人首先提出。他们意识到语境本身的信息不足以理解他人的意图，人们利用面部表情资源来分析他人的想法和愿望。该版本RMET包括25张男女演员展示眼周面部区域的照片。受试者被要求在两个词中选择出最能描述照片中人物想法或感觉的词语，这些词语包括描写基本心理状态（如快乐）和复杂心理状态（如傲慢）的词语。旨在通过使用眼周表情评估参与者识别他人心理状态的能力，来评估成人的社会认知。

由于项目的数量和两个选项的格式并没有充分区分获得较高分数的个体，该版本有一定的局限性。因此，Baron-Cohen等在2001年对RMET进行修订，照片数量由25张增加到36张，每个项目的可选项由2项增加到4项（描述可能心境的单词），最大正确猜测率减少到25%。可能的心理状态的描述主要为复杂心境。该测试有足够的分析复杂性以适用于有或无精神疾病、脑损伤或痴呆的成年人，可评估有平均智力的成年人群的社会认知。修订后的RMET在社会敏感性方面提高了检测细微个体差异的能力。

RMET不仅是先进的心理理论测试，也用于评估情感识别。完成该测试不仅需要识别情绪表达的能力，还需要基于个体部分面部表情判定复杂的认知心理状态的能力。在拥有这些能力的前提下，人们才能识别一个心理状态的词汇，并了解心理状态术语的意义。

一、不同语言版本

多项社会认知障碍的临床研究表明，典型个体的RMET得分显著高于精神分裂症、孤独症、进食障碍和社交焦虑患者，这些研究表明，RMET可用于评估成年人的社会认知。RMET也被证明可用于评估不同文化背景下的社会智力及其轻度损伤。RMET已经被翻译成多国语言，包括中文、土耳其语、意大利语、匈牙利语、瑞典语、希腊语、日语、法语、德语、阿根廷语和西班牙语等。

1. RMET-西班牙语版　RMET-西班牙语版同英语原版一样，受试者被展示不同个体的36张眼区图片，被要求为每张图片在4个选项中选择出最能描述该图片心理状态的词语。每回答正确记1分，满分36分。

2. RMET-巴西版　RMET-巴西版最初的版本有37个项目，修订后的RMET-巴西版有36个项目外加练习项目，保留原英文版本格式。

3. RMET-中文版 RMET-中文版选用的36个项目、各项目中的4个选项及正确答案均与Baron Cohen等的2001版RMET相同，仅将照片更换成中国人群的眼区图像。测试后，将正常组完成率<50%，且在正常组、MCI组和AD组的比较中均无显著差异的项目剔除，组成RMET-S(RMET-short)，共20项。

二、指导语与评分

RMET既可以使用纸笔，也可以使用电脑版完成。该测试很简单，检查者不局限于专业人士。RMET英文修订版包含的36张图片均为黑白色的面部相同区域的(从鼻梁中间到眉毛上部)照片，外加1张练习图片，所有照片都同等大小(15 cm×6 cm)。每一张图片周围围绕4个描述心理状态的词语，选项均为复杂的心理状态(如羞耻、内疚、好奇等)。检查者告知受试者，在做出选择之前，请确保已经阅读了所有的4个单词，应该尽可能快地回答，但尽可能不要犯错误。受试者必须选择正确描述图片中表达心理状态的词语。该测试还包括一个词汇表，它是一个包含所有测试单词、每个单词相应的同义词和它们在短语中使用例子的列表。词汇表必须在测试开始前展示给受试者，以防止语言偏差。

在电脑版本中，检查者应告知受试者在20 s内选择最合适的项目，一旦受试者做出了决定，下一张照片就会出现。如果没有在20 s内决定，它会自动进行下一张照片。

每回答正确1张照片记1分，满分36分。RMET结果常常是用总成绩来解释的，然而，一些研究已经分离出因子分。最常见的划分由积极的、消极的和中性的情感项目组成，但这个结果尚未得到公认。

三、影响因素

研究表明，性别对RMET得分有影响(女性优势)，这种性别差异可能部分反映了大脑结构和功能的发展差异，这是在胎儿睾酮和成人缩宫素的影响下形成的，这种影响通过执行功能(如行为抑制和注意力)起作用。此外，职业、移情、艺术观、教育程度等均会影响RMET的表现。也有研究表明，教育程度、研究领域和社会期望对RMET的表现无显著影响。

四、信度、效度与常模

Baron-Cohen等评估RMET修订版的正常值为26.2±3.6分。美国

RMET 相关研究对 25 名健康人群(女性 60%,年龄 26.4±12.0 岁,教育年限 14.8±1.4 年)进行了 RMET 修订版测试,RMET 平均得分为 25.0±3.63 分;另一项美国研究(272 名健康人群,女性 48%)评估 RMET 修订版的正常值为 24.9±5.2 分。英国 RMET 的相关研究表明,女性 RMET 的正常值为 27.4±3.4 分(共 168 人,年龄 39.2±10.8 岁),男性 RMET 的正常值为 25.5±4.6分(共 152 人,年龄 37.7±12.6 岁)。

西班牙的一项研究对 358 名心理学本科生采用 RMET-西班牙语版进行测试,其中女性 283 名,初次测试的平均年龄为 34.2±9.0 岁(18~65 岁),结果 RMET-西班牙语版的平均得分为27.2±3.6 分,1 年后重复测试,得分为 27.2±3.7 分,重测信度为 0.6 ($P<0.01$)。土耳其的一项研究对 117 名健康志愿者进行 RMET-土耳其语版测试(由 36 张不同个人的眼睛区域的照片组成)后,正确项目的平均值为 24.5±3.4 分,重复测试后得分为 24.1±4.4 分。笔者对 95 名(教育年限 12.7±3.0 年,年龄 65.7±8.4 岁,女性 52 名)上海社区健康老人进行 RMET-中文版测试,平均得分为 23.7±4.7 分。这个得分明显低于欧美国家的得分,所以笔者采集更典型的图片,重新进行验证。

有研究对 200 名(男性占 46%)18~32 岁的大学生进行 RMET-意大利语版调查,发现该测试的内部一致性为 0.6,重测信度为 0.8(95%可信区间 0.8~0.9)。RMET-波斯语版在线测试在 545 名 16~69 岁成年人(女性 51.7%)中进行,发现该测验的内部一致性为 0.4,重测信度为 0.7(95%可信区间 0.5~0.9)。

RMET 的不足之处:RMET 按照“全或无”的方式来评价受试者的心理理论能力,没有将心理理论视为一种连续的变量进行评估,而且此任务是离线(off-line)的,受试者并没有参与到实际的社会互动中进行心理理论推理,影响了测验的生态效度。

五、临床应用

RMET 能够测试一个更深入的社会推理的独立领域,而不仅仅是识别面部表情的心理状态的能力。RMET 也是测量个体大脑功能的好方法,最重要的是,其他心理理论测试大多专为评估儿童发育、脑损伤或孤独症设计,而 RMET 在一般成人中有实质性差异,解释其结果要考虑语言与非语言两个方面。RMET 已被广泛应用于精神疾病与神经系统退行性疾病的研究。

Morris 等发现眼区和整个面部一样能唤起人类杏仁核活动的增加。

RMET 表现差与左侧额下回病变相关，左侧额下回是理解眼区心理的关键区域，很可能涉及语义工作记忆系统的更普遍的损伤，该系统便于通过眼睛的表达来推理别人的感觉和思想。在青春期前，人类从他人眼睛中解读感觉和想法的能力已产生。在所有年龄段，阅读眼区心理时可在其后颞上沟发现大脑活动。只有早期的青少年在检测时表现出内侧前额叶皮质、额中回和颞极的额外参与。

（马　洁　郭起浩）

第二节　失言认知测验(FPT)

失言认知测验（faux pas test，FPT）由 Valerie Stone 和 Simon Baron-Cohen 等于 1999 年编制，由 10 个失言故事与 10 个没有失言的对照组故事组成。所谓失言，就是某人无意之中说的话引起了他人尴尬。笔者翻译了的 FPT 的内容如下，同道在应用时可以把英文的姓名、地址改为中国人习惯的名称。笔者的使用体会是 FPT 对社会认知障碍有一定的敏感性，但检测耗时偏长。

一、操作

印一份只有故事没有问题的版本放在患者面前。对他说："我将念几个短篇故事，并问你一些关于故事的问题。你前面放了一份故事的复印件，因此你可以跟着读或返回阅读。"

接着朗读这些故事，并问下面的问题。如果患者对第一个问题的回答是"没有，没有人说了什么不该说或显得尴尬的话"，直接跳到此故事的控制问题。无论患者对于"有没有人说了尴尬的话?"回答"有"或"没有"，请确保一定要问控制问题。

二、评分

对于含有一个失言之处的故事（故事 2、4、7、11～16、18），患者每答对一个问题可得 1 分。

问题 1："是否有人说了什么不该说或显得很尴尬的话?"

失言问题　　正确：是；错误：否

控制问题 错误：是；正确：否

问题 2:“是谁说了不该说或显得尴尬的话呢?”

任何可以明确表示正确的人的答案皆可以算正确。

关于将小女生叫成小男生的故事：Mary(可接受：邻居)

关于水晶碗的故事：Anne(可接受：屋主、结婚的女人等)

关于律师的故事：Claire(可接受：那个女人、心情不好的女人等)

关于窗帘的故事：Lisa(可接受：那位朋友)

关于癌症玩笑的故事：Jake(可接受：迟到的那个人)

关于输掉故事比赛的故事：Jake(可接受：赢的男生)

关于打翻咖啡的故事：Tim(可接受：打翻咖啡的人)

关于学校新学生的故事：Joe(可接受：Joe 和 Peter)

关于惊喜派对的故事：Sarah(可接受：打翻咖啡的女人)

关于派的故事：Joe(可接受：Kim 的表兄)

对于问题 1 回答“否”的患者不会被问这题，并且这题记为 0 分。

问题 3:“为什么他/她不该说那些话，或者为什么会显得尴尬?”

任何合理的、可以解释故事中失言之处的答案都可以被接受。

患者无须特别提到精神状态，如“他不知道有人得了癌症，但其他人都知道”。说出以下回答已足够:“因为 John 已经得了绝症”或“因为站在那里的男人娶了一位律师”或“你不应该走进一个新公寓又批评它，你不知道是谁买了什么”。

如果患者的答案无法反映对失言之处(也就是可能显得冒犯之处)的理解，这个问题只被记为不正确。

举例:“Claire 不应该说她需要她的咖啡”(忽略对 Roger 的冒犯)“你参加会议不应该迟到”(没有提到开得不适当的玩笑)。

对于问题 1 回答“否”的患者不会被问这题，并且这题记为 0 分。

问题 4:“他们为何那样说?”或“你认为他/她为什么会说那些话?”

同样地，任何合理的、可以解释故事中失言之处的答案都可以被接受。只要患者的回答显示他理解故事中的某个角色并不知道某件事或没有发觉某件事都算正确，就算没有特别提到精神状态。如果患者认为那个失言的人是故意说出了那些话，这个问题被记做错误。

对于问题 1 回答“否”的患者不会被问这题，并且这题记为 0 分。

问题 5:“X 知道 Y……吗?”

同样地，这是要测试他们是否知道失言是无意的。打分很简洁明了。

问题6:"X有何感受?"

一个测试患者对于故事中角色的同理心的问题。应该反映出受伤、愤怒、尴尬、失望的感受,是合理的。

问题7和问题8:控制问题。

这些问题可以告诉你患者是否搞混并忘记故事的细节了。答案都很显而易见。这些问题与其他问题应分开评分。

失言故事的举例:

"故事中,Sally在哪里?""在她阿姨Carol家。"

"故事中,Jeanette送了Anne什么结婚礼物?""一个水晶碗。""一个碗。"

"故事中,Robert的太太是做什么职业的?""她是一名律师。"

"故事中,Jill刚买了什么?""新的窗帘。""窗帘。"

"故事中,主管Jean在会议中跟员工说了什么?""VP得了癌症。"

"故事中,谁赢了比赛?""Jake。"

"故事中,Jack站在哪里?""收银机旁边。"

"故事中,Joe与Peter在谈话时Mike在哪里?""在厕所间里。"

"故事中,Helen的先生举办惊喜派对是为谁办的呢?""Helen。"

"故事中,Kim做了什么口味的派?""苹果。"

举例来说,额叶背外侧受损的患者经常将这些弄错。一位患者说:"惊喜派对为Sarah的生日办的,Helen之所以生气是因为她的丈夫为另一个女人举办派对,且她担心他们是不是有婚外情。"

所有的患者都需要被问这些问题,就算他对于问题1回答了"否"。

总的来说,患者在10个失言故事中对问题的回答一共能得到总分60分。当患者对问题1回答了"否",他整个故事将会获得0分。

在10个控制故事中,患者如果正确地说出没有人说了什么不该说的话可得2分,如果回答有人说了不该说的话则得0分。控制故事的总分一共是20分。每答对一题控制问题则得1分。

分开记录失言故事的失言相关问题得分、失言故事的控制问题得分、控制故事的失言相关问题得分和控制故事的控制问题相关得分。之后你能够有个整体概念,患者是否在回答失言相关问题(心理理论错误)时犯的错误较回答控制问题时多。

如果有患者将任何控制问题回答错误,那么他们在此故事中的其他错误应当被小心地解释。你可以排除患者在此故事中的其他回答,并只对剩余的

故事进行评分，将总分改成54分或48分等进行答对率的评分。

如果对问题1和问题5的回答不符，应该做记录。

为了减少篇幅，笔者对评分标准中各种错误回答的大量举例基本删除了。

附9-1　失言认知测验(FPT)的故事材料与提问

1. Vicky在她朋友Olivia家参加一个派对。当她正在跟Olivia说话时，Olivia的其中一位邻居上前与她们打招呼，她说完"你好！"又转向Vicky说："我想我们还没见过面。我是Maria，请问你叫什么名字？""我叫Vicky。"

Olivia问："有没有人想喝点饮料呢？"

问题：

"是否有人说了什么不该说或显得很尴尬的话？"

如果有，请问：

"是谁说了不该说或显得尴尬的话呢？"

"为什么他/她不该说那些话或为什么会显得尴尬？"

"你认为他/她为什么会说那些话？"

"你认为Olivia知道Vicky和Maria并不认识吗？"

"你认为Vicky有何感受呢？"

控制问题：

"故事中，Vicky在哪里？"

"Vicky和Maria认识彼此吗？"

2. Helen的丈夫打算为Helen办一场生日惊喜派对。他邀请了Helen的一位朋友Sarah，并对她说："不要告诉任何人，尤其是Helen。"生日派对的前一天，Helen在Sarah家，Sarah不小心将一些咖啡打翻在挂在椅背上的洋装上。"哦！"Sarah说："我原本打算穿这件去参加你的派对的！""什么派对啊？"Helen问，Sahara接着说："走吧！我们去看看能不能将这个污渍去掉。"

问题：

"是否有人说了什么不该说或显得很尴尬的话？"

如果有，请问：

"是谁说了不该说或显得尴尬的话呢？"

"为什么他/她不该说那些话或为什么会显得尴尬？"

"你认为他/她为什么会说那些话？"

"Sarah记得这个派对是一个惊喜派对吗？"

"你认为 Helen 有何感受呢?"

控制问题:

"故事中,惊喜派对是为谁举办的?"

"是什么被打翻到洋装上了?"

3. Jim 去逛街想买一件可以搭配他套装的衬衫。店员给他看了好几件衬衫,Jim 看完终于找到一件他想要的颜色的衬衫,但是当他到试衣间试衣服时才发现衬衫的大小不合适。Tim 对店员说:"我想这件太小了。""别担心,我们下周还会进更大尺码的货。"店员回应道。Jim 接着说:"太好了,那我到时再回来。"

问题:

"是否有人说了什么不该说或显得很尴尬的话?"

如果有,请问:

"是谁说了不该说或显得尴尬的话呢?"

"为什么他/她不该说那些话或为什么会显得尴尬?"

"你认为他/她为什么会说那些话?"

"当 Jim 试衣服时,他知道这件衬衫没有他的尺码吗?"

"你认为 Jim 有何感受呢?"

控制问题:

"故事中,Jim 要买什么东西?"

"为什么他两周后要再回店里?"

4. Jill 刚搬进一间新公寓。她去逛街并为她的卧室买了一些新窗帘。当她刚布置好公寓,她最好的朋友 Lisa 前来拜访。Jill 带 Lisa 参观了她的公寓后问道:"你喜欢我的卧室吗?"Lisa 说:"那些窗帘真是太糟糕了!我希望你会换新的!"

问题:

"是否有人说了什么不该说或显得很尴尬的话?"

如果有,请问:

"是谁说了不该说或显得尴尬的话呢?"

"为什么他/她不该说那些话或为什么会显得尴尬?"

"你认为他/她为什么会说那些话?"

"Lisa 知道那些窗帘是谁买的吗?"

"你认为 Jill 有何感受?"

控制问题：

“故事中，Jill刚买了什么东西？”

“Jill在这间公寓住了多久？”

5. Bob到理发店里剪头发。理发师问他：“你想要剪成什么样子？”Bob说：“我想要跟现在一样的造型，只是帮我剪短一寸左右。”理发师将前侧的头发剪得有些不对称，因此他必须再剪短才能修饰对称。理发师说：“抱歉，可能剪得比你原本想要的短。”Bob回应：“那也没办法，它会再长出来的。”

问题：

“是否有人说了什么不该说或显得很尴尬的话？”

如果有，请问：

“是谁说了不该说或显得尴尬的话呢？”

“为什么他/她不该说那些话或为什么会显得尴尬？”

“你认为他/她为什么会说那些话？”

“当Bob在剪头发时，他知道理发师会剪得太短吗？”

“你认为Bob有何感受？”

控制问题：

“故事中，Bob想要将头发剪成什么样子？”

“理发师将他的头发剪成了什么样子？”

6. John在回家途中停靠加油站为他的车子加油。他将信用卡交给收银员，收银员将它刷过柜台上的刷卡机。收银员说：“对不起，这台机器无法识别你的信用卡。”John说：“嗯，真奇怪！那我就用现金付吧。”John给了收银员20元并说：“我要加满无铅汽油。”

问题：

“是否有人说了什么不该说或显得很尴尬的话？”

如果有，请问：

“是谁说了不该说或显得尴尬的话呢？”

“为什么他/她不该说那些话或为什么会显得尴尬？”

“你认为他/她为什么会说那些话？”

“当John将信用卡交给收银员时，他知道机器无法识别他的卡吗？”

“你认为John有何感受？”

控制问题：

“故事中，John停下车来买什么？”

“为什么John要用现金付款?”

7. Sally是一个3岁的小女孩,有着一张圆脸和金色短发。她在她阿姨Carol家时,门铃响了,Carol去开了门。来的是一位邻居Mary,Carol说:“嗨,你过来拜访真好。”Mary说:“你好。”并对着Sally说:“哦,我想我还没见过这个小男孩。你叫什么名字?”

问题:

“是否有人说了什么不该说或显得很尴尬的话?”

如果有,请问:

“是谁说了不该说或显得尴尬的话呢?”

“为什么他/她不该说那些话或为什么会显得尴尬?”

“你认为他/她为什么会说那些话?”

“Mary知道Sally是个女孩吗?”

“你认为Sally有何感受?”

控制问题:

“故事中,Sally在哪里?”

“是谁来拜访了?”

8. Joan带着他的狗Zack到公园。她丢了一支棍棒给Zack捡。他们在公园溜达了一阵子后,Joan的邻居Pam经过并跟她聊了几分钟。Pam问:“你正要回家吗?想要一起走吗?”Joan回应:“当然好。”她呼叫Zack,但它正忙着追赶鸽子没有回来。Joan说:“看来它还没准备好要走,我想我们再待一会儿。”Pam回应:“好的,那我们待会儿见。”

问题:

“是否有人说了什么不该说或显得很尴尬的话?”

如果有,请问:

“是谁说了不该说或显得尴尬的话呢?”

“为什么他/她不该说那些话或为什么会显得尴尬?”

“你认为他/她为什么会说那些话?”

“当Pam邀请Joan时,她知道Joan无法跟她一起回家吗?”

“你认为Pam有何感受?”

控制问题:

“故事中,Joan带着Zack去了哪里?”

“为什么她没有跟她的朋友Pam一起走?”

9. Joanne 在去年的校剧中扮演了重要的角色,今年她真的很想担任主角。为此她学习了戏剧课程,到了春天她参加了校剧的试镜。结果公布的那天,她特地去了学校,查看角色当选名单。她不但没有被选上主角,反而只担任一个小配角。她在走廊上遇到了她的男朋友,并对他说发生了什么事。她男朋友说:“我很抱歉,你一定感到很失望。”Joanne 接着说:“是的,我还要决定是否要演出这个角色。”

问题:

“是否有人说了什么不该说或显得很尴尬的话?”

如果有,请问:

“是谁说了不该说或显得尴尬的话呢?”

“为什么他/她不该说那些话或为什么会显得尴尬?”

“你认为他/她为什么会说那些话?”

“当 Joanne 的男朋友刚在走廊上遇到她时,他知道她没有被选上主角吗?”

“你认为 Joanne 有何感受?”

控制问题:

“故事中,Joanne 得到了什么角色?”

“Joanne 在前一年担任的是什么样的角色?”

10. Joe 在图书馆里,他找到了他想要的关于在大峡谷登山的那本书,并前往柜台办理出借。当他查看皮夹时,才发现他把图书证忘在家里了。他对柜台的女士说:“对不起,我好像把图书证忘在家里了。”女士回答:“没关系,你只需要告诉我你的名字,只要我们的电脑系统中能查到,你就可以凭驾照租借这本书。”

问题:

“是否有人说了什么不该说或显得很尴尬的话?”

如果有,请问:

“是谁说了不该说或显得尴尬的话呢?”

“为什么他/她不该说那些话或为什么会显得尴尬?”

“你认为他/她为什么会说那些话?”

“Joe 进入图书馆时,他发现自己没有带图书证吗?”

“你认为 Joe 有何感受?”

控制问题:

“故事中,Joe 在图书馆借了什么书?”

"他有办法租借这本书吗?"

11. Jean 是 Abco 软件设计公司的一位主管,她组织了一场员工大会。会上她说:"我有一件事要告诉大家,我们的一位会计 John Morehouse 得了癌症,身体非常虚弱,正在医院治疗。"现场一片寂静,大家都在消化这个噩耗,这时,一位工程师 Robert 迟到了,他说:"嘿!我昨晚听到一个很棒的笑话!你们猜猜一位绝症患者跟他的医生说了什么?"Jean 说:"好的,我们在会议中继续谈公事吧。"

问题:

"是否有人说了什么不该说或显得很尴尬的话?"

如果有,请问:

"是谁说了不该说或显得尴尬的话呢?"

"为什么他/她不该说那些话或为什么会显得尴尬?"

"你认为他/她为什么会说那些话?"

"当 Robert 进来时,他知道那位会计得了癌症吗?"

"你认为主管 Jean 有何感受呢?"

控制问题:

"故事中,主管 Jean 在会议中跟员工说了什么?"

"是谁在会议中迟到了?"

12. Mike 是一个 9 岁的小男孩,刚转入一所新的学校。他正在学校厕所的一隔间内,这时另外两位男生 Joe 和 Peter 进来并站在水槽边聊天。Joe 说:"你知道班上那个新同学吗?他叫 Mike。你不觉得他很奇怪吗?而且他好矮哦!"Mike 走出厕所隔间,Joe 和 Peter 看到了他。Peter 说:"哦,你好啊,Mike!你现在要去外面踢足球吗?"

问题:

"是否有人说了什么不该说或显得很尴尬的话?"

如果有,请问:

"是谁说了不该说或显得尴尬的话呢?"

"为什么他/她不该说那些话或为什么会显得尴尬?"

"你认为他/她为什么会说那些话?"

"当 Joe 在和 Peter 说话时,他知道 Mike 在其中的一个厕所隔间吗?"

"你认为 Mike 有何感想?"

控制问题：

“故事中，Joe 和 Peter 在讲话时 Mike 在哪里？”

“Joe 说了什么关于 Mike 的事情？”

13. Kim 的表兄 Scott 要前来拜访，Kim 为了他特别做了一个苹果派。吃完晚饭后，Kim 说：“我特地为你做了一个派在厨房里。”Scott 回应：“嗯，真好闻，我最爱吃派了，当然除掉苹果的。”

问题：

“是否有人说了什么不该说或显得很尴尬的话？”

如果有，请问：

“是谁说了不该说或显得尴尬的话呢？”

“为什么他/她不该说那些话或为什么会显得尴尬？”

“你认为他/她为什么会说那些话？”

“当 Scott 闻到派时，他知道那是苹果口味的吗？”

“你认为 Kim 有何感想？”

控制问题：

“故事中，Kim 做了什么样的派？”

“Kim 和 Scott 是怎么认识彼此的？”

14. Jeanette 为他的朋友 Anne 买了一个水晶碗作为结婚礼物。Anne 举办了一场盛大的婚礼，并且要记录的礼物很多。大约一年后，有一天晚上 Jeanette 到 Anne 家吃晚饭，她不小心将一个酒瓶砸到水晶碗上并将碗打碎了。Jeanette 说：“我真抱歉，把碗打碎了。”Anne 回应：“别担心，反正我从来就不喜欢那个碗，它是别人送我的结婚礼物。”

问题：

“是否有人说了什么不该说或显得很尴尬的话？”

如果有，请问：

“是谁说了不该说或显得尴尬的话呢？”

“为什么他/她不该说那些话或为什么会显得尴尬？”

“你认为他/她为什么会说那些话？”

“Anne 还记得那个碗是 Jeanette 送的吗？”

“你认为 Jeanette 有何感受？”

控制问题：

“故事中，Jeanette 送了 Anne 什么生日礼物？”

“碗是怎么被打碎的?”

15. 在Fernhaven小学举办了一场说故事比赛,所有人都被邀请。好几位五年级学生参加了。其中一位五年级学生Christine非常喜爱她参加比赛的故事。过了几天,比赛的结果公布了,Christine的故事并没有获得任何奖项,她班上的同学Jake获得了头奖。隔天Christine与Jake坐在长凳上一起看Jake的头奖奖杯。Jake说:“这个比赛赢得真容易,其他参赛的故事都好糟糕。”Christine接着问道:“你打算把奖杯放在哪里呢?”

问题:

“是否有人说了什么不该说或显得很尴尬的话?”

如果有,请问:

“是谁说了不该说或显得尴尬的话呢?”

“为什么他/她不该说那些话或为什么会显得尴尬?”

“你认为他/她为什么会说那些话?”

“Jake知道Christine写了故事参加比赛吗?”

控制问题:

“故事中,谁赢了这场比赛?”

“Christine的故事赢了任何奖项吗?”

16. Tim在一家餐厅,不小心将咖啡打翻在地。服务员说:“我再帮你倒一杯咖啡。”并走开了一阵子。Jack是餐厅里的另一位顾客,正站在收银台准备结账。Tim走向Jack,对他说:“我打翻了一些咖啡在桌子旁,能帮我拖一下吗?”

问题:

“是否有人说了什么不该说或显得很尴尬的话?”

如果有,请问:

“是谁说了不该说或显得尴尬的话呢?”

“为什么他/她不该说那些话或为什么会显得尴尬?”

“你认为他/她为什么会说那些话?”

“Tim知道Jack是另一位顾客吗?”

“你认为Jack有何感想?”

控制问题:

“故事中,Jack为什么站在收银台旁边?”

“Tim打翻了什么?”

17. Eleanor 在车站等公交车。由于公交车晚点，她已在那里站了好长一段时间。她已经 65 岁了，站立这么久让她感到劳累。当公交车终于来了，上面已经非常拥挤，且没有剩余的座位了。她看到一位邻居 Paul 站在公交车的走道上，Paul 对她说："你好，Eleanor，你在那里等了很久吗？" Eleanor 回应，"大约 20 分钟。"之后一位坐在座位上的年轻人站了起来说："阿姨，你要坐吗？"

问题：

"是否有人说了什么不该说或显得很尴尬的话？"

如果有，请问：

"是谁说了不该说或显得尴尬的话呢？"

"为什么他/她不该说那些话或为什么会显得尴尬？"

"你认为他/她为什么会说那些话？"

"当 Eleanor 上公交车时，Paul 知道她已经等了好久吗？"

"你认为 Eleanor 有何感受？"

控制问题：

"故事中，为什么 Eleanor 在公交车站等了 20 分钟？"

"她上公交车时车上还有剩余的空位吗？"

18. Roger 刚开始在一家新公司工作。一天，他在茶水间跟他的新朋友 Andrew 聊天。Andrew 问："你太太是做什么职业的？""她是一位律师。"Roger 回答。过了几分钟，Claire 一脸厌烦地进入茶水间，她告诉他们："我刚接了一个最痛苦的电话，律师都好自大又贪婪，我真受不了他们！"Andrew 接着问 Claire："你要过来过目这些报告吗？"Claire 回答："现在不要，我需要我的咖啡。"

问题：

"是否有人说了什么不该说或显得很尴尬的话？"

如果有，请问：

"是谁说了不该说或显得尴尬的话呢？"

"为什么他/她不该说那些话或为什么会显得尴尬？"

"你认为他/她为什么会说那些话？"

"Claire 知道 Roger 的太太是一名律师吗？"

"你认为 Roger 有何感受？"

控制问题：

"故事中，Roger 的太太是做什么职业的？"

“Roger 和 Andrew 在哪里聊天？”

19. Richard 买了一台红色的标志新车。几个星期后，Richard 倒车时撞到他邻居 Ted 的老沃尔沃车。他的新车没有任何损伤，且也没有对 Ted 的车造成大伤害，只有在轮胎上方有一道小刮痕。尽管不严重，他仍然去敲了 Ted 的大门，当 Ted 应门时，Richard 说：“真的很抱歉，我在你的车上撞了一道小刮痕。”Ted 看了看车子说：“别担心，这只是个意外。”

问题：

“是否有人说了什么不该说或显得很尴尬的话？”

如果有，请问：

“是谁说了不该说或显得尴尬的话呢？”

“为什么他/她不该说那些话或为什么会显得尴尬？”

“你认为他/她为什么会说那些话？”

“Richard 知道 Ted 会有什么反应吗？”

“你认为 Ted 有何感受？”

控制问题：

“故事中，Richard 对 Ted 的车子做了什么？”

“Ted 对此有何反应？”

20. Louise 到屠宰店买一些肉。店里既拥挤又吵闹，她问了屠夫：“你们还有散养鸡吗？”他点了点头并开始为她包装一只烤鸡。Louise 说：“不好意思，我想我可能没有说清楚，我是问你们还有散养鸡吗？”屠夫回答：“哦，对不起，我们都卖完了。”

问题：

“是否有人说了什么不该说或显得很尴尬的话？”

如果有，请问：

“是谁说了不该说或显得尴尬的话呢？”

“为什么他/她不该说那些话或为什么会显得尴尬？”

“你认为他/她为什么会说那些话？”

“当屠夫开始为 Louise 包装烤鸡时，他知道她要的是一只散养鸡吗？”

“你认为 Louise 有何感想？”

控制问题：

“故事中，Louise 去了哪里？”

“为什么屠夫开始包装一只烤鸡给她?”

（郭起浩）

第三节　爱荷华博弈任务(IGT)

爱荷华博弈任务(Iowa gambling task, IGT)由 Bechara 等模拟真实生活中的复杂决策情景设计出来，用于验证“脑损伤患者不能利用过去经验对将来的事件做出决策”的观点。该任务是测量情感决策功能的经典实验任务，其最大特点是模仿现实生活中的决策情景：既有奖励，也有惩罚，且奖惩具有不确定性。之后该试验范式被广泛用于与决策有关的研究。相关研究认为，IGF 不仅包括了决策的前提和结论，而且包括了奖励-损失及其数量和频率的多重变化；同时 IGF 也是一项涉及情绪、记忆与学习、认知评价、奖赏和运动程序编制与执行等多个系统协同活动的决策任务。

一、最初的爱荷华博弈任务版本(Bechara 等，1994)

在 1994 年 Bechara 等设计的爱荷华博弈任务中，受试者需要在立即获得一个大奖励(100 美元)而长远看来不利的两副纸牌(A 和 B；每 10 张纸牌损失 250 美元)和立即获得一个小奖励(50 美元)但长远看来有利的两副纸牌(C 和 D；每 10 张纸牌赢得 250 美元)中进行选择(表 9－3－1)。不利牌的每一张比有利牌提供更高的赢钱数(100 ∶ 50)，但输钱数不固定且不可预测；有利牌虽每次赢钱数较少，但损失是固定的，且总体来说远小于不利牌。所以连续选择有利牌最后会净赢，而持续选择不利牌则会导致净输。测试中每次给受试者呈现 4 张牌，每一张牌翻开后，都揭示不同的赢钱和输钱数目。检查者需记录受试者每次选择的纸牌类型。例如：第一次受试者在 B 中选择，检查者需在 B 的第一格记录“1”；第二次受试者还是在 B 中选择，检查者需在 B 的第二格记录“2”；第三次受试者在 A 中选择，检查者需在 A 的第一格记录“3”，如此选满 100 次。若受试者选中了输钱的纸牌，检查者就告诉他：“这次你赢了 100 或 50 美元，但你同时输掉 X 美元。”

指导语：“你面前有 A、B、C、D 4 副纸牌，每次从任意一副纸牌中选 1 张，然后交给我，我会展示给你看。每一张牌翻开后你都会赢得一定数目的钱，但我现在不会告诉你将会赢得多少钱，测试中你会慢慢明白；有时你会输钱给

我，我现在也不会告诉你何时输钱及输多少，测试中你也会慢慢明白。每次你可以从任意一副纸牌中选择 1 张，目的是尽可能多赢，尽量避免输钱。你需要一直选择下去，直到我说停为止。需要提醒你的是，纸牌的颜色与本测试无关，你无法从颜色判断赢钱或输钱，有几副牌好于另外几副。请把纸牌上的钱当作你自己的钱认真对待，现在我给你 2 000 美元，结束时我会收回，并计算你赢钱或输钱的数目。纸牌上的'0'与空牌相同。"

表 9-3-1　最初的爱荷华博弈任务版本的奖励-损失结构(Bechara 等，1994)

IGT	A(+100)	B(+100)	C(+50)	D(+50)
1		0		0
2		0		
3	−150		−50	0
4		0		
5	−300		−50	
6		0		0
7	−200		−50	0
8				
9	−250	−1 250	−50	
10	−350		−50	−250
11		0		0
12	−350	0	−25	0
13			−75	
14	−250	−1 250		
15	−200		−25	0
16		0		
17	−300	0	−25	
18	−150		−75	
19				0
20			−50	−250
21		−1 250		
22	−300			0
23		0		
24	−350	0	−50	
25			−25	0
26	−200	0	−50	0
27	−250			0

（续表）

IGT	A(+100)	B(+100)	C(+50)	D(+50)
28	−150	0		
29			−75	−250
30			−50	
31	−350	0		
32	−200	−1 250		0
33	−250			
34		0	−25	0
35			−25	−250
36		0		0
37	−150		−75	
38	−300	0		
39			−50	0
40			−75	
净值（40 次中）	−1 000	−1 000	1 000	1 000
奖励-损失频率（每 10 次中）	10 次奖励，5 次损失	10 次奖励，1 次损失	10 次奖励，5 次损失	10 次奖励，1 次损失

二、最初的爱荷华博弈任务的变异版本(Bechara 等,2000)

2000 年，Bechara 等设计了爱荷华博弈任务的变异版本，与原始版本基本相同，仅奖惩顺序与原始版本相反（表 9-3-2）。变异版本包括 E、F、G、H 4 副纸牌，受试者需要在立即遭遇一次大损失（100 美元）而长远看来有利的两副纸牌（E 和 G；每 10 张纸牌赢得 250 美元）和立即遭遇一次小损失（50 美元）但长远看来不利的两副纸牌（F 和 H；每 10 张纸牌损失 250 美元）中进行选择。若受试者选到了赢钱的纸牌，检查者就告诉他："这次你输了 100 或 50 美元，但你同时赢得了 X 美元。"

指导语与 1994 年版本类似。

表 9-3-2　最初的爱荷华博弈任务变异版本的奖励-损失结构(Bechara 等,2000)

IGT	E(−100)	F(−50)	G(−100)	H(−50)
1	0			0
2			350	
3	1 250			0

（续表）

IGT	E(−100)	F(−50)	G(−100)	H(−50)
4	0	25	250	
5		50		
6	0		300	0
7			200	0
8	0	75		250
9		25	150	
10		75		
11	1 250	50		0
12				0
13		25	350	
14	0			
15			250	0
16	0	25		
17	0	75	200	
18			150	
19				0
20	0	75	300	250
21	1 250			
22			300	
23	0			
24		25	350	
25	0	75		0
26		50	150	0
27	0		200	0
28	0		250	
29		75		0
30		25		250
31	0		150	
32			200	
33	1 250		350	0
34	0	50		250
35		50		
36	0			0
37	0	25	200	
38			350	0

（续表）

IGT	E(−100)	F(−50)	G(−100)	H(−50)
39		75		0
40		50		
净值(40 次中)	1 000	−1 000	1 000	−1 000
奖励-损失频率(每 10 次中)	10 次损失，1 次奖励	10 次损失，5 次奖励	10 次损失，5 次奖励	10 次损失，1 次奖励

三、临床爱荷华博弈任务版本(Bechara, 2007)

临床版本与最初版本的基本原理相同，存在以下不同之处。第一，每次的奖励不同，A 和 B 两副纸牌每次奖励 80～170 美元，平均 100 美元，C 和 D 两副纸牌每次奖励 40～95 美元，平均 50 美元。第二，每 10 张纸牌奖励金额的平均数递增（A 和 B 递增 10 美元，C 和 D 递增 5 美元）。第三，A 和 C 两副纸牌，每 10 张纸牌中输钱牌的数量递增，而输钱数一致；B 和 D 两副纸牌，每 10 张牌中输钱牌的数量相同，而输钱数递增。如 A(C)前 10 张牌中有 5 张不利纸牌，每张输钱数 150～350(25～75)美元，第二个 10 张牌中有 6 张不利纸牌，每张输钱数也为 150～350(25～75)美元；而 B(D)前 10 张牌中有 1 张不利纸牌，输钱数 1 250(250)美元，第二个 10 张牌中也只有 1 张不利纸牌，而输钱数增加至 1 500(275)美元。第四，A 和 B 两副纸牌中每 10 张纸牌的输钱数递增 150 美元，而 C 和 D 两副纸牌中每 10 张纸牌的赢钱数递增 25 美元。

指导语与之前的版本类似。

表 9-3-3　临床爱荷华博弈任务版本的奖励-损失结构(Bechara, 2007)

IGT	A	B	C	D
每次奖励	80～170	80～170	40～95	40～95
1				
2				
3	−150		−50	
4				
5	−300		−50	
6				
7	−200		−50	
8				
9	−250	−1 250	−50	

（续表）

IGT	A	B	C	D
10	−350		−50	−250
11				
12	−350		−25	
13			−75	
14	−250	−1 500		
15	−200		−25	
16				
17	−300		−25	
18	−150		−75	
19	−250			
20			−50	−275
21	−250	−1 750		
22	−300		−25	
23				
24	−350		−50	
25			−25	
26	−200		−50	
27	−250			
28	−150		−25	
29	−250		−75	−300
30			−50	
31	−350		−25	
32	−200	−2 000		
33	−250		−25	
34	−250		−25	
35	−150		−25	−325
36				
37	−150		−75	
38	−300		−25	
39	−350		−50	
40			−75	
41	−350		−25	
42	−200			
43	−250		−25	
44	−250		−25	

（续表）

IGT	A	B	C	D
45	−150		−25	−350
46		−2 250	−25	
47	−150		−75	
48	−300		−25	
49	−350		−50	
50	−250		−75	
51	−350		−25	
52	−200		−25	
53	−250		−25	
54	−250		−25	
55	−150		−25	
56	−250		−25	
57	−150		−75	
58	−300	−2 500	−25	−375
59	−350		−50	
60	−250		−75	
净值(60 次中)	−3 750	−3 750	1 875	1 875
奖励-损失频率(每 10 次中)	10 次奖励，5～10 次损失	10 次奖励，1 次损失	10 次奖励，5～10 次损失	10 次奖励，1 次损失

Kerr 和 Zelazo 简化了 Bechara 等研究中的爱荷华博弈任务，从而设计了儿童博弈任务(children's gambling task, CGT)。此任务运用了两副纸牌，一副纸牌的正面是竖条花纹，另一副的正面是圆点花纹。将两副纸牌翻过来都能看见它们的反面有开心的脸和悲哀的脸。不同的是，正面是竖条纹的纸牌的反面总是有 1 张开心的脸，偶尔加上 1 张悲哀的脸；而正面是圆点花纹的纸牌的反面总是有 2 张开心的脸，但有时会出现好几张(如 4、5、6 张不等)悲哀的脸。开心的脸代表赢得糖果，其数量也代表赢得糖果的数量；悲哀的脸代表输掉糖果，其数量也同样代表输掉糖果的数量。每次试验只能选取一张纸牌。显然，选竖条花纹的纸牌虽然每次赢的糖果更少，只有 1 颗，但平均起来，输的糖果也更少；相反，选圆点花纹的纸牌虽然每次赢的糖果更多，有 2 颗，但平均损失却大得多，一旦输，就会输掉 4 颗、5 颗或 6 颗。因此，从长远来看，选竖条花纹的纸牌有利；反之，则不利。试验中，研究者告诉儿童"游戏"结束时要赢得尽量多的糖果(如 50 次选牌，这点儿童事先不知)。开始的 25 次选择可以

看作是儿童对两种纸牌的尝试；后面的 25 次试验将被作为对情感决策的诊断。此试验的关键因变量为儿童在第 26～50 次试验中做出不利选择的比例。这种研究方法主要是对儿童情感决策的诊断，通过这种诊断可以判断儿童的“热”执行功能发展与年龄有关的规律。它还考察儿童控制受即时愿望支配的动作的能力和预测他们的动作将会产生后果的能力。它常用于学龄前儿童，但对学龄儿童也适用。他们测查了三四岁儿童在任务中的表现，结果发现 3 岁幼儿无法通过该任务。

Garon 和 Moore 仍使用 4 副牌，试验次数为 40 次，研究了 3 岁、4 岁和 6 岁儿童在任务中的表现，结果发现即使是 6 岁的儿童也只处于概率水平，但是他们随后用 CGT 范式研究 3.5 岁和 4.5 岁的儿童时得出了和 Kerr 等相似的结论。然而 Bunch 等在 CGT 的基础上继续简化任务的复杂性，将之前 Kerr 等研究中涉及的牌、收益、损失三者之间的三元关系修改为收益或损失（另一个保持恒定）与牌的二元关系，发现 3～5 岁的儿童均能通过该任务。

四、评分

常用的有 3 种记分方法：① 方法一：(C+D)−(A+B)或(E+G)−(F+H)，该记分方法可以分离受试者的冒险策略。该净值>0，表示受试者选择冒险；反之，净值<0，表示受试者选择保守。② 方法二：分为五个 Block，每个 Block 含 20 次选择，(C1+D1)−(A1+B1)……(C5+D5)−(A5+B5)，然后做配对 T 检验，该记分方法可以分离随机与思考之间的决策成绩——归纳学习。③ 方法三：低分组和高分组，以总分 20 分为界，该记分可以区分不同的受试者或人群。

五、评价

爱荷华博弈任务的设计源自对腹内侧前额皮质受损（ventromedial prefrontal cortex，VM-PC）患者的观察和研究。在 Bechara 等的研究中，通过对不同受试群体的研究结果显示，正常健康个体对纸牌的选择会随着试验的进行逐渐倾向于有利，相反，腹内侧前额皮质受损的个体倾向于做冒险的选择。他们的表现类似于成瘾者、精神变态者以及一些自毁者的行为：这些患者坚持那些从长远来看对他们的健康、人际关系以及经济稳定等方面带来不利后果的行为，可以将其称为“短见主义者(myopia for the future)”。之后该任务被广泛应用，2007 年 Bechara 等设计的版本可用于评价多种神经精神疾

病所致的决策障碍，包括局灶性脑损伤、吸毒成瘾、病理性赌博、精神分裂症、强迫症、神经性厌食症、肥胖、慢性疼痛、多动症、攻击性疾病、情感障碍及Huntington舞蹈病等。有这些神经精神疾病的患者往往较多选择不利的两副纸牌，而正常人倾向于选择有利的两副纸牌。

爱荷华博弈任务既为人们模拟出一种更接近现实生活的决策情景，也为研究者探究某些患者群体的心理特点提供了可能。利用爱荷华博弈任务对患者群体的研究发现，它能很好地测查出患者个体在现实生活中的决策缺陷，而且任务中某些参数的变化并不影响成绩。比如，在爱荷华博弈任务中无论是用真币还是假币，无论对个体的反应时间做何要求，无论是用手工版本还是用计算机版本，个体的任务成绩都是非常稳定的。因此，它被研究者认为是测量个体模糊决策能力的经典实验室任务。

爱荷华博弈任务的信度和效度受多种因素影响。有研究者认为，受试者个体化差异和情感状态可影响效度，另有研究认为“B纸牌优势（PDB）”现象（指一部分正常人更喜欢不利的B纸牌，赢钱频率高但期望值为负值）会对效度产生影响。因此，该研究的信度和效度有待在更广泛的正常人群中进行考察获取。

（王　萍　郭起浩）

第十章

行为症状量表

第一节　神经精神量表(NPI)

神经精神量表(neuropsychiatric inventory questionnaire,NPI)是Cumming等于1994年针对痴呆患者所呈现的精神病理改变而设计的。本书所附版本是笔者根据英文版本翻译的,可能与国内其他单位采用的版本在内容与语句上有一些差别(表10-1-1)。神经精神量表-记录表(NPI)(表10-1-2)是根据医生询问知情者或家属的信息填写的。神经精神问卷—临床医生等级评定量表(neuropsychiatric inventory-clinician rating scale, NPI-C)还有医生对患者的面谈评估,由于版权关系无法附录在本书里。

指导语:(面对照料者)对于以下行为是否存在,请您回答"是"或"否"。如果相符请在后面的方框里打钩。并请给这些行为的发生频率(1～4)和严重度(1～3)分别评分。两者的乘积(A×B)为患者的该项得分。再给知情者的痛苦程度评分(0～5)。

表10-1-1　神经精神量表完全版

评定者:		日期:
患者个人信息:		
姓名:	年龄:	性别:M/F
住院/疗养院/门诊		

请您回答"是"或"否"对于以下行为是否存在或者这些筛检问题并没有提到。如果相符请在后面的方框里打钩。并请给这些行为的发生频率(1～4)和严重度(1～3)分别评分。两者的乘积(A×B)为患者的该项得分。再给您的痛苦程度评分(0～5)。

（续表）

一、妄想

该患者是否存在某种信念，但其内容与事实不符？例如，坚信有人要毒害或窃听他/她？他是否不承认家里的亲人或者房子？我并非出于猜疑这么问，我只是想知道患者是否坚信这些事情在他身上发生？

不适用的　否（跳到下一个筛检问题）。

是（回答以下问题）。

1. 他认为有人计划要害他而使他正处于危险中？
2. 他认为被人监视或窃听？
3. 他认为他的配偶有外遇？
4. 他认为有不受欢迎的客人住在自己家中？
5. 他认为他的配偶或其他人不是他们所宣称的那个人？
6. 他认为他的家实际不是他的家？
7. 他认为亲人计划着离弃他？
8. 他认为电视或杂志里的人物实际上存在家里？（他是否尝试和他们对话或彼此影响？）
9. 他相信有某种我不曾提到的异常的事物？

A. *发生频率指数*：

偶尔—每周少于一次。 1
经常—每周一次。 2
频繁—每周几次，但非每天。 3
非常频繁—每天一次或更多。 4

B. *严重度指数*：

轻度—存在妄想，但看似并非有害，且患者本人很少痛苦。 1
中等—妄想使人苦恼。 2
显著—妄想非常严重以至于成为行为异常的主要来源（如果有药物干预，本身就代表了妄想的严重性）。 3

C. *痛苦度指数*：

你认为该项行为对你所造成的情感上的苦恼有多少？
根本没有。 0
略微有一点。 1
轻度。 2
中等。 3
严重。 4
极度严重。 5

该项得分（A×B）：

二、幻觉

该患者是否有幻视或幻听诸如此类的错觉？他是否看到、听到或经历某些事实上并不存在的事情？这个问题想问的并不是某人已经死了而你却坚信他还活着，而是该患者是否真听到或看到不寻常的事物？

不适用的　否（跳到下一个筛检问题）。

是（回答以下问题）。

1. 他有幻听或有幻听的表现？
2. 他和某个不存在的人对话？
3. 他有幻视或有幻视的表现（诸如看到人、动物、光等）？
4. 他有幻嗅？

（续表）

5. 他感觉有东西爬在或触摸他的皮肤或者表现出这种感觉？
6. 他描述某种味道但找不到任何原因？
7. 他描述另一些不同寻常的感知体验？

A. *发生频率指数*：

偶尔—每周少于一次。	1
经常—每周一次。	2
频繁—每周几次，但非每天。	3
非常频繁—每天一次或更多。	4

B. *严重度指数*：

轻度—存在幻觉，但看似并非有害，且患者本人很少痛苦。	1
中等—幻觉使人苦恼。	2
显著—幻觉非常严重以至于成为行为异常的主要来源（如果有药物干预，本身就代表了幻觉的严重性）。	3

C. *痛苦度指数*：

你认为该项行为对你所造成的情感上的苦恼有多少？

根本没有。	0
略微有一点。	1
轻度。	2
中等。	3
严重。	4
极度严重。	5

该项得分（A×B）：

三、激越

该患者是否拒绝合作或者不想别人帮助他？他是否很难对付？

不适用的　否（跳到下一个筛检问题）。

是（回答以下问题）。

1. 他会对那些想要帮助他的人产生不安或是拒绝，像洗澡、更衣这样的活动？
2. 他很顽固，凡事必须按他自己的方式进行？
3. 他不合作，拒绝从别人那里得到帮助？
4. 他有一些别的行为从而使得他很难相处？
5. 他咒骂或是大发雷霆？
6. 他撞门，踢家具，乱摔东西？
7. 他设法伤害或攻击别人？
8. 他有一些攻击或激越的行为？

A. *发生频率指数*：

偶尔—每周少于一次。	1
经常—每周一次。	2
频繁—每周几次，但非每天。	3
非常频繁—每天一次或更多。	4

B. *严重度指数*：

轻度—存在激越，但看似并非有害，且患者本人很少痛苦。	1
中等—激越使人苦恼。	2
显著—激越非常严重以至于成为行为异常的主要来源（如果有药物干预，本身就代表了激越的严重性）。	3

C. *痛苦度指数*：

你认为该项行为对你所造成的情感上的苦恼有多少？

根本没有。	0

(续表)

略微有一点。	1
轻度。	2
中等。	3
严重。	4
极度严重。	5

该项得分(A×B):

四、抑郁/心境恶劣

该患者是不是看上去悲伤或抑郁的？或者他说自己感到悲伤或抑郁？

不适用的　否(跳到下一个筛检问题)。

是(回答以下问题)。

1. 他流泪或呜咽,好像很悲伤?
2. 他说出或表现出好似很悲伤,情绪低落?
3. 他贬低自己,感觉自己很失败?
4. 他说自己是个糟糕的人物,应该受到惩罚?
5. 他看上去很气馁,说自己没有未来可言?
6. 他说自己对于家庭是一个负担,家庭没有他的存在会更好?
7. 他有想死的念头,甚至自杀?
8. 他有另一些抑郁或悲伤的表现?

A. 发生频率指数:

偶尔—每周少于一次。	1
经常—每周一次。	2
频繁—每周几次,但非每天。	3
非常频繁—每天一次或更多。	4

B. 严重度指数:

轻度—存在抑郁,但看似并非有害,且患者本人很少痛苦。	1
中等—抑郁使人苦恼。	2
显著—抑郁非常严重以至于成为行为异常的主要来源(如果有药物干预,本身就代表了抑郁的严重性)。	3

C. 痛苦度指数:

你认为该项行为对你所造成的情感上的苦恼有多少?

根本没有。	0
略微有一点。	1
轻度。	2
中等。	3
严重。	4
极度严重。	5

该项得分(A×B):

五、焦虑

该患者是否没有缘由的非常紧张,担心,害怕？看上去紧绷,烦躁的？害怕远离你？

不适用的　否(跳到下一个筛检问题)。

是(回答以下问题)。

1. 他说他担心明明已经计划好的事情?
2. 他感觉摇摆不定,不能放松或感觉非常紧张?
3. 他没有理由地呼吸短促,喘息或叹息?

（续表）

4. 他抱怨怀里像揣着一只兔子忐忑不安的(排除生病状态)？
5. 他避免某些特定的会使他紧张的场面如驾驶车，和朋友见面或在人群里？
6. 一和你(或他的关照人)分开他会变得很紧张或不安(他黏住你不让你离开他)？
7. 他有另一些焦虑的表现？

A. *发生频率指数*：

偶尔—每周少于一次。	1
经常—每周一次。	2
频繁—每周几次，但非每天。	3
非常频繁—每天一次或更多。	4

B. *严重度指数*：

轻度—存在焦虑，但看似并非有害，且患者本人很少痛苦。	1
中等—焦虑使人苦恼。	2
显著—焦虑非常严重以至于成为行为异常的主要来源(如果有药物干预，本身就代表了焦虑的严重性)。	3

C. *痛苦度指数*：

你认为该项行为对你所造成的情感上的苦恼有多少？

根本没有。	0
略微有一点。	1
轻度。	2
中等。	3
严重。	4
极度严重。	5

该项得分(A×B)：

六、情感高涨/欣快

该患者是不是看上去异常兴奋却毫无理由的？这里我指的不是正常人的高兴，比如说看到朋友，收到礼物或是和家人在一起。我指的是是否该患者有一种持久的不寻常的高涨情绪。

不适用的　否(跳到下一个筛检问题)。

是(回答以下问题)。

1. 他看上去过分高兴，和他以往不同？
2. 他发现某种乐趣和嘲笑某种事物而别人却不觉得有趣？
3. 他表现出幼稚，会吃吃地笑，甚至不适事宜地大笑(例如某人发生了不幸的事)？
4. 他喜欢讲笑话或说一些别人并不觉得有趣的话，但似乎自己觉得很有趣？
5. 他喜欢玩幼稚的恶作剧并自得其乐？
6. 他说大话，宣称自己既有钱又有才，然而并不真实？
7. 他有其他一些自我感觉很好或很欣快的表现？

A. *发生频率指数*：

偶尔—每周少于一次。	1
经常—每周一次。	2
频繁—每周几次，但非每天。	3
非常频繁—每天一次或更多。	4

B. *严重度指数*：

轻度—存在欣快，但看似并非有害，且患者本人很少痛苦。	1
中等—欣快使人苦恼。	2
显著—欣快非常严重以至于成为行为异常的主要来源(如果有药物干预，本身就代表了欣快的严重性)。	3

（续表）

C. 痛苦度指数：
你认为该项行为对你所造成的情感上的苦恼有多少？
根本没有。 0
略微有一点。 1
轻度。 2
中等。 3
严重。 4
极度严重。 5

该项得分(A×B)：

七、情感淡漠

该患者是否对他周围的世界失去了兴趣？他是否失去了做事的积极性或缺少开创新活动的动力？他是不是很难投入会话或做家务杂事？他是不是冷漠的？

不适用的 否(跳到下一个筛检问题)。

是(回答以下问题)。
1. 他似乎没有以往积极和主动？
2. 他参与对话的可能性比较小？
3. 若和以往的自己相比现在他缺乏感情？
4. 他现在做家务活比以前少了？
5. 他似乎对别人的活动或计划感兴趣的少了？
6. 他丧失了对朋友和家人的关心？
7. 他对自己以往的爱好也少热衷了？
8. 他有其他一些他不在乎新事物的表现？

A. 发生频率指数：
偶尔—每周少于一次。 1
经常—每周一次。 2
频繁—每周几次，但非每天。 3
非常频繁—每天一次或更多。 4

B. 严重度指数：
轻度—存在淡漠，但看似并非有害，且患者本人很少痛苦。 1
中等—淡漠使人苦恼。 2
显著—淡漠非常严重，以至于成为行为异常的主要来源(如果有药物干预，本身就代表了淡漠的严重性)。 3

C. 痛苦度指数：
你认为该项行为对你所造成的情感上的苦恼有多少？
根本没有。 0
略微有一点。 1
轻度。 2
中等。 3
严重。 4
极度严重。 5

该项得分(A×B)：

八、去抑制

他似乎不经思考，表现冲动？他做或说一些公众场合不太做或说的事情？他做一些让你或他人尴尬的事情？

不适用的 否(跳到下一个筛检问题)。

（续表）

是（回答以下问题）。
1. 他不经思考，莽撞行事？
2. 他和陌生人说话好像认识一样？
3. 他对人们说一些伤害他们感情的事情？
4. 他会说一些以前从来不说的脏话或摆出性感的姿态？
5. 他开放地谈论一些在公众场合不宜谈论的私人话题？
6. 他行为自由，和别人有过分亲密动作？
7. 他有另一些失去控制的表现？
A. 发生频率指数：
偶尔—每周少于一次。 1
经常—每周一次。 2
频繁—每周几次，但非每天。 3
非常频繁—每天一次或更多。 4
B. 严重度指数：
轻度—存在淡漠，但看似并非有害，且患者本人很少痛苦。 1
中等—淡漠使人苦恼。 2
显著—淡漠非常严重以至于成为行为异常的主要来源（如果有药物干预，本身就代表了淡漠的严重性）。 3
C. 痛苦度指数：
你认为该项行为对你所造成的情感上的苦恼有多少？
根本没有。 0
略微有一点。 1
轻度。 2
中等。 3
严重。 4
极度严重。 5

该项得分（A×B）：

九、易激惹/不稳定

他容易生气？他的情绪很多变？性情急躁？我指的不是挫折失意或无能完成某项任务，我指的是他是否有不同寻常的易激惹，不耐烦或多变无常的情绪而以往不曾有？

不适用的　否（跳到下一个筛检问题）。

是（回答以下问题）。
1. 他有一个坏脾气，为一点小事就动怒？
2. 他的情绪变化无常，这一分钟还很好，下一分钟就生气？
3. 他会突然发怒？
4. 他没有耐心，没能力处理日常琐事或等待计划好的活动？
5. 他性情古怪，易激惹？
6. 他爱争论，很难相处？
7. 他有另外一些易激惹的表现？
A. 发生频率指数：
偶尔—每周少于一次。 1
经常—每周一次。 2
频繁—每周几次，但非每天。 3
非常频繁—每天一次或更多。 4
B. 严重度指数：
轻度—存在激惹，但看似并非有害，且患者本人很少痛苦。 1
中等—激惹使人苦恼。 2

（续表）

显著—激惹非常严重以至于成为行为异常的主要来源（如果有药物干预，本身就代表了激惹的严重性）。 3

C. 痛苦度指数：

你认为该项行为对你所造成的情感上的苦恼有多少？

根本没有。 0

略微有一点。 1

轻度。 2

中等。 3

严重。 4

极度严重。 5

该项得分（A×B）：

十、迷乱的动作行为

他是否步调缓慢，反复做同样的事情，比如开橱或抽屉？或者反复缠绕线团？

不适用的　否（跳到下一个筛检问题）。

是（回答以下问题）。

1. 他在房间里无目的地踱步？
2. 他一遍遍地搜查衣橱和抽屉？
3. 他重复穿上和脱下衣服？
4. 他有重复的行为或习惯使得他一遍又一遍地去做？
5. 他重复一些行为比如扣纽扣，包裹线团？
6. 他非常坐立不安，似乎不能安静坐下来，上蹿下跳，不停敲打手指？
7. 他有另外一些重复的动作？

A. 发生频率指数：

偶尔—每周少于一次。 1

经常—每周一次。 2

频繁—每周几次，但非每天。 3

非常频繁—每天一次或更多。 4

B. 严重度指数：

轻度—存在，但看似并非有害，且患者本人很少痛苦。 1

中等—使人苦恼。 2

显著—非常严重以至于成为行为异常的主要来源（如果有药物干预，本身就代表了严重性）。 3

C. 痛苦度指数：

你认为该项行为对你所造成的情感上的苦恼有多少？

根本没有。 0

略微有一点。 1

轻度。 2

中等。 3

严重。 4

极度严重。 5

该项得分（A×B）：

十一、睡眠

他是否睡眠不好（一晚爬起来1～2次去盥洗间回来又立刻入睡的不算在内）？他晚上兴致很高？他在夜晚徘徊，打扮，扰你入眠？

不适用的　否（跳到下一个筛检问题）。

是（回答以下问题）。

1. 他入睡很困难？

（续表）

2. 他晚上中途会起床（如果他是起来1～2次去盥洗间，然后回来立刻入睡的不计在内）？
3. 他在夜晚徘徊，踱步，做不适事宜的活动？
4. 他在晚上会把你弄醒？
5. 他在晚上把你弄醒，穿上衣服然后计划外出，并认为已经是早晨了可以开始新的一天？
6. 他在早晨醒得很早（比他以往的习惯早）？
7. 他白天睡得很多？
8. 他有另一些会打扰你的晚间活动而我不曾提起？

A. *发生频率指数：*

偶尔—每周少于一次。	1
经常—每周一次。	2
频繁—每周几次，但非每天。	3
非常频繁—每天一次或更多。	4

B. *严重度指数：*

轻度—存在，但看似并非有害，且患者本人很少痛苦。	1
中等—使人苦恼。	2
显著—非常严重以至于成为行为异常的主要来源（如果有药物干预，本身就代表了严重性）。	3

C. *痛苦度指数：*

你认为该项行为对你所造成的情感上的苦恼有多少？

根本没有。	0
略微有一点。	1
轻度。	2
中等。	3
严重。	4
极度严重。	5

该项得分（A×B）：

十二、食欲和饮食障碍

他在食欲、体重和饮食习惯上有所改变（被喂进去的也算在内）？他是否在事物的类别选择上也有所改变？

不适用的　否（跳到下一个筛检问题）。

是（回答以下问题）。
1. 他食欲不振？
2. 他食欲大增？
3. 他体重有减？
4. 他体重增长？
5. 他在饮食习惯上有所改变，比如一次放很多食物在嘴巴里？
6. 他喜欢的食物种类有改变，比如吃很多糖果或别的特别的食物类别？
7. 他改变了自己的饮食习惯，比如每天吃几乎同一种类别的食物或进食几乎按同样的顺序？
8. 他有另一些食欲或饮食上的改变而我不曾提到？

A. *发生频率指数：*

偶尔—每周少于一次。	1
经常—每周一次。	2
频繁—每周几次，但非每天。	3
非常频繁—每天一次或更多。	4

B. *严重度指数：*

轻度—存在，但看似并非有害，且患者本人很少痛苦。	1
中等—使人苦恼。	2
显著—非常严重以至于成为行为异常的主要来源（如果有药物干预，本身就代表了严重性）。	3

（续表）

C. 痛苦度指数：	
你认为该项行为对你所造成的情感上的苦恼有多少？	
根本没有。	0
略微有一点。	1
轻度。	2
中等。	3
严重。	4
极度严重。	5
该项得分(A×B)：	

表 10-1-2　神经精神量表(NPI)记录纸

内　容	发生频率	严重度	痛苦度
	1 2 3 4	1 2 3	0 1 2 3 4 5
1. 妄想(delusions)			
2. 幻觉(hallucinations)			
3. 激越(agitation)			
4. 抑郁/心境恶劣(depression)			
5. 焦虑(anxiety)			
6. 情感高涨/欣快(euphoria)			
7. 情感淡漠(apathy)			
8. 去抑制(disinhibition)			
9. 易激惹/不稳定(irritability)			
10. 迷乱的动作行为(aberrant motor behavior)			
11. 睡眠(night-time behaviors)			
12. 食欲和饮食障碍(appetite and eating disorders)			

第二节　曼彻斯特和牛津大学痴呆病理心理学评定量表(MOUSEPAD)

曼彻斯特和牛津大学痴呆病理心理学评定量表(Manchester and Oxford Universities scale for the psychopathological assessment of dementia, MOUSEPAD)是笔者根据 Allen 等 1996 年发表(Br J Psychiatry)的版本翻译的(表 10-2-1)。它的优点是包括一些罕见的行为症状，如错觉、饮食异常等。

表 10-2-1　曼彻斯特和牛津大学痴呆病理心理学评定量表(MOUSEPAD)完全版

评定者：　　　　　　　　　　　　　　　　　　　日期：
患者个人信息：
姓名：　　　　　　　　　　年龄：　　　　　　　性别：M/F
住院/疗养院/门诊

首先，确定一下痴呆综合征持续的时间是以月来计的：多久前你第一次注意到这个症状？(以下为了方便起见我们都用“他”或“她”代替)

这个评定量表可以评定当前正发生的以及从痴呆一开始就存在的症状和行为。0＝不存在；1＝轻度，症状明显的不常发生(一周少于1次)；2＝中度，症状发生得较频繁(一周1次或更多，但少于一周4次)；3＝严重，症状发生至少一周4次；8＝不适用的或没有提及；9＝被访问者不知道。

一、妄想

这些信念很坚定，但据查证是错误的，持续时间要大于7天，并且不是发生在有急性器质性疾病的基础上。即使这些错误的信念可能和记忆损害有关，但还是应该被评定。请看下面。

他是否曾经说过：	是/否	严重度	几月(从开始)	持续时间(月)	上个月？是/否	能被说服？	得分
他被监视？							
他的事物或饮料被下毒？							
他被跟踪？							
他的附体被隐藏了？							
他的附体被偷走了？							
他的房子不再是他自己的房子？							
他的配偶有外遇？							
他被人暗恋或钟情？							
他被家人所遗弃？							
有别人在自己房子里？							

二、幻觉

如果这个患者自发地抱怨某种现象或者有证据表明他被错误的观念所影响，那这种体验就被认为是存在的。这样的体验必须发生在没有急性器质性疾病的基础上，并且已经持续了大于7天时间。请注意：患者和镜子里的自己、照片、电视说话并不算在内，这些应该归在错误识别里。

他是否听到其实并不存在的声音？如果是，这些声音是：	是/否	严重度	几月(从开始)	持续时间(月)	上个月？是/否	得分
嗓音						
来自熟知的人						
来自陌生人						
他看上去理解这种声音吗？						
音乐						
动物						
别的响声						

（续表）

他是否看到其实并不存在的东西？如果是，这个东西是：	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
人						
熟知的人						
陌生人						
侏儒						
孩子						
动物						
别的						

他是否曾经说过：	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
他是否说过闻到其实不存在的奇怪的气味？						
他是否说过有种奇怪的感觉在身上？						
他是否说过他的食物或饮料里有种不寻常的味道？						

三、错误识别

这些体验必须持续至少7天，并且发生在没有急性器质性疾病的基础上。

当他看到镜中的自己，反映强烈？如果是，他：	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
宣称在镜中的形象非他自己？						
花时间和镜中的自己说话？						
他是否明确相信：						
他的亲戚或照顾者不是他们自己所宣称的那个人？						
他们被一个骗子所替代？						
电视或照片上的形象是真实的事件？						
他被小动物所侵扰？						

四、复制

这些体验必须持续至少7天，并且发生在没有急性器质性疾病的基础上。

他是否说过东西被复制了（例如，一个人变两个人或任何别的东西）？如果是，是什么东西？	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
配偶或照顾者？						
房子或别的没有生命的事物？						
有生命的事物（例如，宠物，等）？						

（续表）

五、痴呆的行为改变

标记下任何可能会更正评定结果的身体状况（例如，关节炎可能会影响他的走路）。如果这个身体状况影响评定结果，请评定为8，并且做一下解释。

六、步行

在记忆问题出现之前，他比以往走路徘徊多了？	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
1. 轻度：大多数日子能15 min里都坐着，并且醒着。						
2. 严重：大多数日子15 min里很难坐着，并且醒着。						
他是否老跟随着你或别人？						
0. 每天跟随别人少于30 min。						
1. 几乎每天跟随别人大于30 min却少于2 h。						
2. 几乎每天跟随别人要大于2 h。						
他是否在房子外面徘徊（超出花园范围）？						
0. 几乎不或偶尔（特殊的日子也少于1 h）。						
1. 大多数日子少于每天3 h。						
2. 大多数日子超出每天3 h。						
他是否在夜晚徘徊（除了他已经睡觉后）？						
他是否徘徊远离家门，要被别人带回来？						

七、进食

起病以来他的体重有否改变？　　没有/明显增长/明显下降

在问题出现之前，他比以往吃得多了？	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
1. 稍微多一点？						
2. 比以前多一倍，或更多？						
在问题出现之前，他吃得比以往快了？						
1. 稍微快一点。						
2. 快很多。						

（续表）

在问题出现之前，他比以往吃得多了？	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
你是否曾限制他进食的量，否则他会吃更多？						
1. 偶尔一些场合会 2. 必须大多数时间都要控制他的摄入						
他比以往吃更多的甜食？						

八、睡眠

他是否出现睡眠问题？	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
他夜晚没有休息的，一直醒着？						
他白昼倒置？						
他白天打盹儿打得比以往问题出现前多了？						

九、性行为

他是否有下列表现？	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
他是否会不适当地谈论性？						
他表现出很性感？						

十、攻击性

他是否有下列表现？	是/否	严重度	几月（从开始）	持续时间（月）	上个月？是/否	得分
自记忆问题出现以来，他是否行为和言辞上都爱寻衅？						
如果是的话，是在哪种环境下？当别人照顾他时（例如，清洗，穿衣）？						
无缘无故的或是不耐烦的？						
对幻觉或是错误地以为你要伤害他的一种反应？						
这种攻击反应在： 行为上—抵触别的人？ 言辞上—抵触事物？						
他会爆发性地： 笑 哭						

（续表）

十一、最近一个月另外的行为表现

他是否有下列行为表现?	是/否	严重度	几月(从开始)	持续时间(月)	上个月?是/否	得分
他把物品藏起来或者放到莫名的地方?						
他丢失了事物?						

（郭起浩）

第三节　额叶损害行为改变评估量表(DEX、FBI)

目前常用的额叶损害行为改变评估量表有5种，执行功能之行为评估清单(behavior rating inventory of executive functions，BRIEF)、执行缺陷问卷(dysexecutive questionnaire，DEX)、额叶行为量表(frontal systems behavior scale，FrSBe)、额叶行为清单(frontal behavior inventory，FBI)和爱荷华人格改变评估量表(Iowa rating scales of personality change，IRSPC)。

本节介绍DEX与FBI。

一、执行缺陷问卷(DEX)

DEX的20个项目用来评估以下特征：抽象思维问题、冲动、交谈、保持、计划问题、欣快、缺少洞察力、淡漠、去抑制、随境转移、知识-反应分离、对社会规则缺少关注或不关注。每一个项目都用5分制打分，从“从不”到“很经常”，内容见表10-3-1。

表10-3-1　执行缺陷问卷(DEX)

注意：首先注明是受试者自评，还是家属或知情者评估，还是评定员评估。

如果是家属或知情者评估，应注意让患者回避。若有知情者或家属不理解的项目，应该给予解释。表格中“当前”：根据近1个月表现的症状严重程度圈出相应选项。

0=无/从不；**1**=轻度/偶尔(1～2天/周)；**2**=中度/经常(3～4天/周)；**3**=重度/大部分时间(5～7天/周)；**4**=全部时间。

序号	项　目	得　分
1	**抽象思维问题**：患者缺乏抽象的逻辑推理与演绎归纳能力，对他人的言语只能从字面去理解。	0 1 2 3 4
2	**冲动**：患者能遵守社会规范吗？他/她是否行为不合规范？比如随意打扰别人的工作、学习和休息，不遵守公共秩序等。患者做事情前不假思索、不顾后果吗？	0 1 2 3 4

（续表）

序号	项　　目	得　分
3	**虚构**：患者是否凭空虚构？没有遇到的事情他/她认为遇到了，并深信不疑。	0 1 2 3 4
4	**计划性**：患者能计划和组织较复杂的活动吗？能够解决工作和生活中遇到的问题吗？保持社会判断吗？	0 1 2 3 4
5	**情感高涨/欣快**：该患者是不是看上去异常兴奋却毫无理由？指的不是正常人应有的高兴，比如看到朋友、收到礼物或和家人在一起的那种高兴。该患者有一种持久的不寻常的高涨情绪吗？	0 1 2 3 4
6	**时间次序问题**：他/她完成一项工作是否缺乏条理或次序？	0 1 2 3 4
7	**自知力缺乏**：患者是否能意识到自己行为、情绪控制、语言、记忆、思维方面的改变？他/她是否倾向于否认这些症状？0＝有自知力，能够正确分析自己病情；1＝承认有病，但不能主动求治（部分自知力）；2＝对自己的症状表现意识到有问题，但是否认是疾病所致；3＝完全否认自己有病。	0 1 2 3 4
8	**淡漠、缺乏动力、（对自己）漠不关心**：患者对自己经历的快乐或悲伤的事情的情感反应是否如常？他/她是否丧失了正常的情感反应？	0 1 2 3 4
9	**失抑制**：表现为过度滑稽行为、失礼行为、漫游行为、胡乱消费行为、收藏行为、收集废品行为、依赖追随行为、偷窃行为、冒险行为。	0 1 2 3 4
10	**反复无常**：想法一会儿这样、一会儿那样，变动不定，做事没有恒心。	0 1 2 3 4
11	**情感反应肤浅**：难以正确理解别人的情感，只能主观臆测；表达自己的情感有困难。	0 1 2 3 4
12	**攻击行为**：攻击行为是指具有对他人有意挑衅、侵犯或对事物有意损毁、破坏等心理倾向和行为。	0 1 2 3 4
13	**兴趣淡漠**：患者对于朋友和日常生活是否不再感兴趣？他/她对于和朋友见面或做事还有兴趣吗？是否对别人的事情不闻不问？	0 1 2 3 4
14	**持续性**：刻板言语或动作。患者是否重复或持续说相同的话或动作？他/她是否机械而刻板地重复某些无意义的言语，持续较长时间，同样的话反复不已；他/她是否持久地、单调地重复一个动作，并没有什么意义和指向？如重复穿上和脱下衣服，反复做同样的事情，比如开橱或抽屉，或者反复缠绕线团。	0 1 2 3 4
15	**多动**：患者很激动、坐立不安吗？	0 1 2 3 4
16	**不能抑制反应**：易怒、易激惹。患者是否很容易发火？缺乏耐心？神经过敏？他/她对于不开心的事情是否反应如常？有无爆发性行为？	0 1 2 3 4
17	**知-行分离**：患者是否明知故犯？	0 1 2 3 4
18	**分心、注意力涣散**：患者能否集中注意力，关注正在发生的事情？他/她是否无法持续跟踪和关注一件事？比如能专心做一件事，不被外界环境所干扰。	0 1 2 3 4
19	**犹豫不决**：患者是否判断力减弱，在做决定的过程中缺乏良好的判断力？	0 1 2 3 4
20	**礼节行为**：患者是否能遵守社会规范？他/她是否行为不合规范？比如随意打扰别人的工作、学习和休息，不遵守公共秩序等。	0 1 2 3 4

DEX 内部一致性很高（＞0.90）。DEX 的评定员之间信度和再测信度尚不明确。

对 DEX 的因子分析表明问卷评估至少会在 3 个或可能更多 5 个领域发

生变化。Wilson 等(1996)报道照料者叙述的症状隐含着 3 个因子,分别代表行为、认知和情感部分,每个因子的唯一有意义预测者是 BADS 得分,认知估计测试、WCST、NART FSIQ 和 WAIS - R FSIQ 表现以及年龄并不是 DEX 表现的预测因素。相比之下,Burgess 等(1998)报道了一个不同组别神经疾病患者的五因子构面(five-factor solution),分别代表抑制、意图、执行记忆、积极作用和消极作用。前 3 个因素与执行任务有良好的相关性,而后面两个的相关性要差很多。Chan(2001)在健康的中国香港人群中发现了相似的五因素构面：抑制、意图、知行背离、抵抗和社会规范,这些因素与其他的标准执行功能测试之间有轻到中度的相关性。Simblett(2012)采用 Rasch 分析区分为 3 个因子：执行认知、元认知、行为-情绪自我调节。

DEX(自评和他评)对于轻到中度 TBI 的存在是敏感的,由专业人士进行的 DEX 评估与损伤的严重程度相关。它是否能区别出额叶病灶还是非额叶病灶仍有待考证。

DEX 自评与他评之间的差异被证明是对自我执行功能缺陷洞察力的有用评估。针对 TBI 患者进行的 DEX 和 BADS 测验发现 DEX 自评报告的问题少于其他人或临床医师所报告的——这就是脑外伤后的一种模式,即对于问题的洞察力下降得很明显。

DEX 得分预测其他执行功能的能力大小存在争议,精神分裂症患者的 BADS 表现和 DEX 评价之间没有明显关系。但是对于脑外伤患者,BADS 总分和 DEX 评价之间关系密切。在其他神经疾病人群中,DEX 他评(DEX - Other)得分与很多神经心理测验中表现出的功能受损相关。相反,DEX 他评/自评得分差异——洞察力的粗略估计,与执行测验得分相关,但是与记忆测验得分和智力水平、命名能力测验无关。Chan(2001)报道轻到中度 TBI 患者的 DEX 他评/自评得分差异与六因素测试表现显著相关,洞察力越差,表现越差,而且与 DEX 自评或他评得分的相关性更高。与汉诺塔测验无明显相关性。但是,也有研究不支持 TBI 患者的两个结论(即 DEX 他评/自评得分差异和他评得分与执行功能相关),这个差异可能由于评定员及评定环境差异所致。Bennett 等发现如果由专业人士(神经心理学家、职业治疗师)来完成的话,通常在一个急性期康复环境中进行 DEX 评定,而由家庭成员进行 DEX 评定,评估执行功能的敏感性就要差一些,患者自评与任何 BADS 测试变量都不相关。

已经有不少证据表明,BADS 具有生态信度,可能是发现日常执行功能障

碍的有用工具。然而，这套测验的研究还是不充分的，比如，没有基于年龄的常模数据，也没有教育程度等影响因素分析的资料，测验的信度偏低，有天花板效应。不同BADS任务之间的相互关系、分测验与总分的关系、与非执行功能测试（如智力水平和空间能力）之间的关系，都是今后值得研究的课题。

DEX提供了一个适用于量化日常执行行为，包括执行缺陷自我意识的评定量表。它经济实用，可以进行自评-他评差异评估，只需要评估员很少的监督。它也适用于非英语环境中（如中国香港）。然而，它的效度证据之间有冲突，部分原因可能是由于评定者不同。有证据表明家庭成员不是总能提供最准确的执行功能评估的，在一些情况下（如急性康复期），长期接触患者（而不是门诊或极短时间交流）的专业人员能够进行更好的评估，优于照料者评估。但是家庭成员比患者更加准确。也就是说，就准确性来说，受试者＜家庭成员或知情者＜了解病情的专业人员。

关于DEX的信度几乎没有公开发表的信息。另外，常模数据也很有限——这是一个很大的限制，因为有证据表明非临床样本在日常生活中也会有执行障碍样的表现。最后，使用者要注意，完成DEX的指导说明是近1个月的表现，若无这个限定时间可能让一些评估者采用一种长期观点来评估患者，而不是当前的行为。

二、额叶行为问卷(FBI)

额叶行为问卷（FBI）对额叶痴呆的诊断准确性最高。它是Kertesz等于1997年编制的结构性问卷，评估行为改变，由有经验的评定员与照料者面谈后填写，不是由患者自评或照料者自行填写的。项目来源于Lund和Manchester与Neary等编制的额颞叶痴呆诊断共识。24个项目见表10-3-2。

表10-3-2　额叶行为问卷(FBI)(询问家属或知情者)

该问卷主要调查患者行为和人格的改变。应单独询问照料者，注意让患者回避。可对问题做必要的解释。根据近1个月表现的症状严重程度圈出相应选项。**0**＝无/从不；**1**＝轻度/偶尔（1～2天/周）；**2**＝中度/经常（3～4天/周）；**3**＝重度/大部分时间（5～7天/周）。

序号	项　目	0 1 2 3
1	**(对外界)兴趣淡漠**：患者对于朋友和日常生活是否不再感兴趣？他/她对于和朋友见面或做事还有兴趣吗？	0 1 2 3
2	**缺乏主动性**：患者能主动做事吗？他/她是否需要别人提出要求才肯去做事？	0 1 2 3
3	**(对自己)漠不关心**：患者对自己经历的快乐或悲伤的事情的情感反应是否如常？他/她是否丧失了正常的情感反应？	0 1 2 3

（续表）

序号	项　　目	0 1 2 3
4	**缺乏灵活性**：患者能合理地改变主意吗？近来他/她是否变得固执、不懂变通？	0 1 2 3
5	**个人忽略**：患者对个人卫生、仪表是否如往常一样在意？他/她是否会忘记换洗内衣？	0 1 2 3
6	**组织能力丧失**：患者能计划和组织较为复杂的活动吗？他/她完成一项工作是否缺乏条理或次序？	0 1 2 3
7	**注意力涣散**：患者能否集中注意力，关注正在发生的事情？他/她是否无法持续跟踪关注一件事？	0 1 2 3
8	**缺乏自知力**：患者是否能意识到自己行为上的改变？他/她是否缺乏自知力或者倾向于否认这些症状？	0 1 2 3
9	**言语减少/找词困难**：患者的言语表达能力是否同以往一样？他/她的言语是否明显减少？	0 1 2 3
10	**语义理解受损**：患者是否难以理解词语及物体，询问词语的含义？他/她是否知道词语的含义？	0 1 2 3
11	**失语和言语失用**：患者是否会用词错误或发音错误？他/她最近是否出现口吃或重复发出某种声音？	0 1 2 3
12	**异己手和(或)失用**：患者是否动作笨拙、手僵硬，无法正常使用物品器具？一手是否会干扰另一手的动作？或表现得好似某只手不属于自己？他/她能如往常一样自如运用双手吗？	0 1 2 3
13	**持续、强迫症状**：患者是否重复或持续做相同的动作或说相同的话？他/她是否有强迫行为或习惯？他/她是否从来都是如此？	0 1 2 3
14	**易激惹**：患者是否很容易发火？他/她对于不开心的事情是否反应如常？	0 1 2 3
15	**过度滑稽**：患者是否不分时间乱开玩笑冒犯他人？他/她一直是个有奇怪的幽默感、乱玩笑的人吗？	0 1 2 3
16	**冲动/判断力减弱**：患者在做决定、花钱、驾驶等过程中是否有良好的判断力？他/她是否缺乏判断力，做事冲动不负责任，不计后果？	0 1 2 3
17	**收藏**：患者是否把东西、钱财过度地收藏？他/她是否以往也一样节俭？	0 1 2 3
18	**行为失范**：患者是否能遵守社会规范？他/她是否说话或行为不规范？是否表现粗鲁或幼稚？	0 1 2 3
19	**坐立不安/漫游**：患者是否过度地行走、踱步、开车？他/她的这些行为还在正常水平吗？	0 1 2 3
20	**激动**：患者是否很激动，对人喊叫或打人？他/她在这方面是否没有变化？	0 1 2 3
21	**食欲亢进**：患者是否食欲亢进，对任何看得见的东西或饮料都要吃喝或放入口中？他/她是否一直胃口很好？	0 1 2 3
22	**性欲亢进**：患者是否性欲亢进或异常（包括脱衣服或涉性言语）？他/她这些方面是否无变化？	0 1 2 3
23	**运用行为**：患者是否对看得到或摸得着的东西都要不自主地触摸把玩一番？他能管住自己的手吗？	0 1 2 3
24	**二便失禁**：患者是否有大小便失禁？他/她的这些问题是否因尿路感染、前列腺病变或生育所致？	0 1 2 3

1. *信度* 评定员之间信度(Cohen κ=0.90)与项目一致性(Cronbach α=0.89)都比较好。

2. *效度* 已经有许多研究发现FBI能够有效鉴别fvFTD与PPA、AD、VaD等，最有鉴别力的项目包括缺乏自知力、淡漠、注意力不集中、持续性、失抑制、行为失范等。推荐划界分27分。目前已经有针对FTLD的药物临床试验采用FBI作为疗效评估工具。

(郭起浩)

第十一章

综合性评估

第一节　阿尔茨海默病评估量表(ADAS)

阿尔茨海默病评估量表(ADAS)包括认知行为(ADAS-cog)与非认知行为量表。认知行为量表包括定向、语言、结构、观念的运用、词语即刻回忆与词语再认,共11题,费时15～30 min,满分70分。非认知量表包括恐惧、抑郁、分心、不合作、妄想、幻觉、步态、运动增加、震颤、食欲改变等10项,每项5分,共50分。根据量表编制者Rosen等报道,对AD组,测试者之间信度为0.99,间隔1个月再测相关性0.92,正常老人组则分别为0.92与0.65。AD患者组在ADAS-cog的每一个项目均显著差于正常匹配组的表现。未经治疗的中度AD患者每年ADAS-cog总分下降7～10分。通常将改善4分(相当于6个月平均自然下降分数)作为临床上抗痴呆药物显效的判断标准。与安慰剂对照组相差2.5分以上才能证明治疗组有效。ADAS是美国FDA批准的目前应用最广泛的抗痴呆药物临床试验的疗效评价工具。

由于ADAS-cog没有详细检测执行功能的项目,在血管性痴呆(VaD)的疗效评定中,修订版的VaDAS-cog增加了范畴流畅性、数字-符号转换、迷宫、数字广度的倒数等测验。ADNI采用的ADAS增加了词语延迟回忆(在图形模仿分测验后进行)和数字划销测验(在一页数字中划去3和7,限时45 s)。Skinner(2012)提出一个扩大版——ADAS-Cog-Plus,是在ADAS-Cog的基础上,增加范畴流畅性、数字-符号转换、连线测验和5项功能评估问卷(FAQ),这个新版本有助于转化为AD的MCI的识别与诊断。

有研究认为从临床药物观察的角度来说,24个月的治疗效果观察,ADNI

采用的13项版本(传统11项+数字消减与延迟回忆)比传统的11项版本更利于出现阳性结果。乙酰胆碱酯酶抑制剂(如多奈哌齐、卡巴拉汀、加兰他敏)治疗阿尔茨海默病主要改善ADAS的记忆项目的得分,近年也不断有研究发现某些治疗措施改善的是ADAS的非常记忆项目的得分,所以,不仅是ADAS总分,其项目分分析也是有必要的。

由于国内翻译的ADAS版本非常多(笔者看到的就有5种),主要差别是第一题词语回忆与第七题词语辨认所采用的词语不一致,有的版本为了治疗前后比较设计了多套难度相似的词语。本书没有提供这两题的词语,一方面是为了避免影响疗效的评估(有些读者购买本书后会事先背这些词语),另一方面是目前的资料无法支持我们选择一个最佳的中文词语版本。ADAS一般在制药公司做临床试验的时候采用,公司会解决版权与版本问题,临床医生很少把它作为临床诊断或鉴别诊断的辅助工具。

阿尔茨海默病评估量表(ADAS)的具体测评步骤和方法介绍如下。

一、初始交谈记录

阿尔茨海默病评估的前10 min进行开放式交谈,以便评估受试者在口语表达和对口语的理解等方面的情况,进行初始交谈记录。然后进行其余的认知测试。语言能力通过整个面谈和特定的测试加以评估。引出"是"或"否"答案的问题,用来评估非常基础层面的理解力。

其他问题需要特定的信息和健全的沟通技巧。鼓励受试者进行简短的交谈,谈论一些中性的话题(例如天气、受试者来诊所途中的情况,或是受试者早餐吃了些什么)。这样的交谈有助于受试者在测试开始前保持放松,并且测试者可以借此观察受试者对语言的运用和理解的情况。在此ADAS的认知部分中,对语言能力有3种临床评级方式。使用表11-1-1来记录你的面谈记录。此份表格中的各种记录应当清晰明确,以便支持口语能力(项目9)、找词困难(项目10)和理解力(项目11)的评级。任何损伤的评级必须得到此表所记录内容的支持。

表11-1-1 阿尔茨海默病初始交谈记录表

可能的话题	食欲	睡眠	锻炼	其他
1. 口语				
2. 找词				
3. 理解力				

二、正式交谈记录

(一) 认知行为

1. 词语回忆(词语列表省略)

(1) 指导语:开始测试时,说:"我将给您出示一些印在白色卡片上的词语。请大声地朗读出每个词语,并且努力记住它。因为稍后,我会要您努力回忆出我给您出示过的所有词语。准备,请读出词语并努力记住它。"

向受试者出示每个词语,并让他/她大声朗读。出示所有10个词语之后,要求受试者尽可能多地回忆词语(词语不需要按照顺序回忆)。

受试者朗读完词语列表后,说:"好了,现在告诉我您能记住的列表中的所有词语。"必要时,提示"还有吗?"在进行第二次和第三次尝试时,说:"现在我要再次给您出示刚才出示过的词语列表。大声读出每个词语并努力记住它。"对于每个被正确回忆的词语,测试者应当勾选"是"。

(2) 评分:对每一个未被回忆起的词语,在"否"方格上打钩,以便计算分数。原版单词包括Coast,Doll,Lip,Chair,Student,Apple,Horse,Valley,Rock,Pipe。计算每次尝试中"否"的方格总数。求尝试1、2和3的评分总和,然后除以3。将得数四舍五入至最近的整数。如果一次尝试未被进行或被受试者拒绝,请不要计算这次尝试的分数。如果一次尝试的分数缺失,请不要记录子分数。

回忆1+2+3/3=即刻词语回忆子分数:(范围=0~10)

2. 执行指令 这个项目用于评估受试者对口语的理解情况。测试者要求受试者执行5个单独的命令,每个命令包括1~5个步骤(表11-1-2)。每个命令只读1次。如果受试者没有反应或看起来没有听懂,或要求重复命令,则测试者应该再次给出完整的命令,然后进入下一个命令。所有命令都要给予每位受试者。测试者应为每一个命令勾选一个回答。

表11-1-2 执行指令执行5个命令记录纸

项 目	对 错
1."握拳"	
2."指向天花板,然后指地板" 在桌子上将铅笔、手表和卡片排成一行(这个顺序是从面对受试者而坐的测试者的角度来看)	
3."把铅笔放在卡片上,然后再把它放回原处"	
4."把手表放在铅笔的另一边,然后把卡片翻过来"	
5."闭上您的双眼,然后用两根手指轻拍您的每侧肩膀两次"	

(1) 指导语：开始测试时，说："现在我要让您做几件事情。首先……"。

(2) 评分：记录执行指令子分数。数一数不正确的回答数量。如果任何任务未进行或被受试者拒绝，不要记录子分数(范围=0～5)。

3. 结构实践　此项测试用于评估受试者临摹4种几何图形的能力(图11-1-1)。图形应每次只出示一个。如果受试者看起来没听懂或对所画图形不满意，或要求再试一次，应允许受试者对每个图形进行第二次尝试。如果进行了第二次尝试，让受试者指出哪次画得更好，并仅对画得更好的那次进行评分。测试者应为每一个图形勾选一个回答(表11-1-3)。

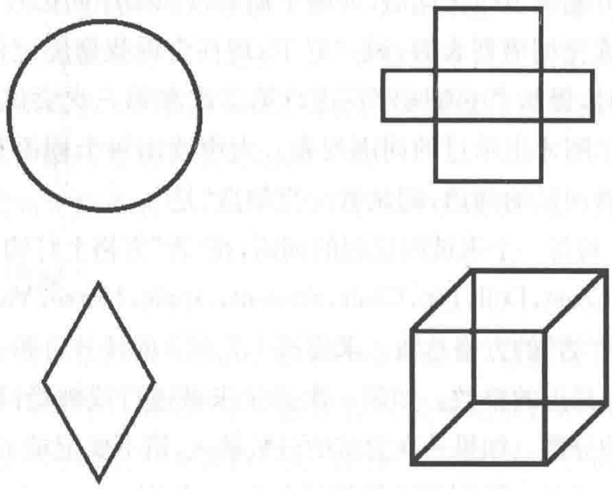


图11-1-1　4种几何图形

表11-1-3　结构实践临摹4种图形记录纸

项　　目	对　错
1. 画圈：一条封闭的曲线	
2. 重叠的两个长方形：必须是两个四边形，而且重叠必须与所示的样子相似 大小的改变不计分	
3. 菱形：必须是四边形，有方向性以使尖端在顶部和底部，而且边要大致等长(例如，最长的边长不能是最短边长的1.5倍或更长)	
4. 立方体：必须是三维立体，内部线条被正确地画在两个角之间。相对的两个面应当大致平行。只要是绘制了正确的立方体，两种方向均可接受	

(1) 指导语：开始测试时，说："在这张纸上画有一个图形。试着在同一张纸上的另一个地方另外画出一个与它非常相似的图形来。"

(2) 评分：评分标准如下：0分为所有4个图形均绘制正确；1分为1个图形绘制不正确；2分为2个图形绘制不正确；3分为3个图形绘制不正确；4分

为4个图形绘制不正确;5为未绘制出图形,乱涂乱画,部分图形。数一数不正确的回答数量,记录结构实践子分数。如果任何图形未被进行或被受试者拒绝,请不要记录子分数(范围=0~5)。

4. 命名物品或手指

(1) 指导语:要求受试者给出表11-1-4中随机出示的12种真实物品的名称。应该按照随机顺序出示物品。给受试者的说明类似于:"现在我将给您出示一些物品。我想要您告诉我它们的名称。这被称为什么?"(出示物品)。如果受试者说出的是物品的用途,则说:"是的,那是这个东西的用途,但是它的名称是什么?"如果受试者不回答,测试者应该给出那件物品的语义提示(在下面提供)。如果受试者还是不回答或出错,则进行下一个物品。测试者应为每一个物品/手指勾选一个回答。然后,要求受试者按表11-1-5中的要求给出他(她)的优势手的手指名称。说:"请将您的右(或左)手放在桌子上。现在我将指出您手上的一个部分,我想让您说出它的名称。这是什么?"

表11-1-4 命令物品测试记录纸

物品名称	语义提示	对 错
花	生长在花园中的	
床	用来睡觉的	
哨子	吹的时候发出声音	
铅笔	用来写字的	
拨浪鼓	一种婴儿的玩具	
面具	遮住你的脸	
剪刀	剪纸的	
梳子	用在头发上的	
钱包	装你的钱的	
口琴	一种乐器	
听诊器	医生用它来听你的心脏	
夹子	夹起食物	

表11-1-5 命令手指测试记录纸

手 指	对 错
拇指/大拇指	
中指	
无名指	
示指(食指)	
小指/小拇指	

（2）评分：将命令物品和命令手指的小计，记录命名子分数。0分为0～2项不正确；1分为3～5项不正确；2分为6～8项不正确；3分为9～11项不正确；4分为12～14项不正确；5分为15～17项不正确。

5. 观念实践　这个项目用于测定受试者能否做出他们熟悉但次序复杂的行动。这个项目有5个任务（表11-1-6）。在受试者面前放一个大信封，一张A4纸和一支铅笔。

表11-1-6　观念实践完成任务记录纸

任　　务	对　错
1. 折纸	
2. 把纸放进信封里	
3. 封好信封	
4. 在信封上收信人的位置写上您自己的姓名和地址	
5. 告诉我邮票贴在哪儿	

（1）指导语："我要您假装给您自己写了一封信。把这张纸拿起来折好，使它与信封的大小相符，然后把它放进信封里。再封好信封，然后在信封上收信人的位置写上您自己的姓名和地址，并且告诉我邮票贴在哪儿。"如果受试者忘记了这项任务的某些部分，或是执行有困难，测试者应当对受试者在这项任务中有困难的部分给予重复说明。给予一次完整的说明后，每项任务只能给予一次额外的提醒。测试者应为每一个任务勾选一个回答。

（2）评分：数一数不正确的回答数量，记录观念实践子分数（范围=0～5）。

6. 定向　这个项目用于测定受试者的时间和空间定向力。

（1）指导语：询问受试者表11-1-7中的信息时应该一次只询问一项。每个问题只允许重新表达一次（例如，当受试者将星期几与日期混淆时）。测试者应为每一个问题勾选一个回答。

表11-1-7　定向测试完成任务记录纸

问　　题	对　错
1. 全名（必须精确）	
2. 月份（必须精确）	
3. 日期（±1天）	
4. 年份（必须精确）	
5. 星期几（必须精确）	
6. 季度（在即将到来的季度1周之内，或在前一个季度的2周之内）	

（续表）

问　　题	对　错
7. 时钟时间(±1小时)(说:“不要看手表,告诉我现在的大概时间。”)	
8. 地点(地名不全是可以接受的,但是不能接受总称)(说:“我们现在在哪里?”或“这个地方叫什么名字?”)	

(2) 评分：数一数不正确的回答数量，记录定向子分数(范围=0～8)。

7. 词语辨认(词语列表省略)

(1) 指导语：在此项测试的记忆部分中，检查者让受试者进行一次尝试，以记忆一份有12个词语的列表(表11-1-8)。说：“我将给您出示一些印在白色卡片上的词语。我要您大声地朗读出每个词语，并且努力记住它。”在此项测试的辨认部分中，测试者应当说：“现在我要给您出示另外一组词语。有些词语在我刚才给您出示的词语列表中出现过，而其他词语则是新的。对于每个词语，我要您告诉我这个词语是否在我刚才给您出示的词语列表中出现过。”测试者出示第一个词语，并说：“这个词语在我以前出示过的词语中吗?是还是不是?”或者说：“我以前向你出示过这个词语吗?”在出示第二个测试词语前，做出同样的说明。对于剩余的测试词语，测试者应该说：“这个词语呢?”可以这样鼓励受试者“尽量猜一下”。勾选受试者对每个词语的回答“对”或“错”。如果受试者在测试期间需要对任务进行提醒，测试者应当重复问题并在提醒栏中做出标记。

表11-1-8　词语辨认测试记录纸

目标词语	对　错	干扰词语	对　错
Silence		Elbow	
Daughter		Powder	
Forehead		Canal	
Tiger		Chamber	
Twilight		Dragon	
Beggar		Sister	
Echo		Nephew	
Village		Duty	
Corner		Music	
Courage		Olive	
Bushel		Ribbon	
Object		Collar	

(2) 评分：数一数已勾选的不正确的方格的数量，记录词语辨认子分数(范围＝0～12)。如果总数等于或少于12，则将总数写在子分数方格中。如果总数大于12，则将12写在子分数方格中。如果任何词语未被进行或被受试者拒绝，请不要计算子分数。

8. **记忆测试说明** 这个项目用来评估受试者对词语辨认任务要求的记忆能力(只能根据在词语辨认任务时的观察)。数一数在词语辨认任务中所给予的提醒次数，以便对该项目进行评级。如果词语辨认任务未被完成或未被尝试，则不得对这个项目进行评分。

损伤水平评分的标准见表11-1-9，记录记忆测试说明子分数(范围＝0～5)。

表11-1-9 记忆测试说明损伤水平评分标准

得分	损伤水平	依据
0	无	受试者从不需要对说明的额外提醒
1	很轻微	忘记1次
2	轻微	必须提醒2次
3	中度	必须提醒3～4次
4	中重度	必须提醒5～6次
5	重度	必须提醒7次或以上

9. **口语能力** 这个项目是对口语质量的全面评级，即清晰度、使自己让人理解的难度。评定这一项时，测试者应当考虑所有受试者在初始面谈和测试环节中说过的话。第一页的各种记录应当清晰明确，以便支持任何0分以上的评级(参考程序手册以获得指南)。

损伤水平评分的标准见表11-1-10，记录口语能力子分数(范围＝0～5)。

表11-1-10 口语能力损伤水平评分标准

得分	损伤水平	依据
0	无	受试者说的话没有让人难以理解的情况
1	很轻微	受试者说的话有一处让人难以理解
2	轻微	受试者说的话让人难以理解的时间少于25%
3	中度	受试者说的话让人难以理解的时间在25%～50%
4	中重度	受试者说的话让人难以理解的时间有50%
5	重度	说出1～2个词；流利，但内容空洞；沉默

10. 找词困难 评估此项时，测试者必须确定在面谈和测试期间，受试者在自发性说话时，找到恰当的词语是否有困难。评估时，不要纳入手指和物品命名。第一页的各种记录应当清晰明确，以便支持任何 0 分以上的评级（参考程序手册以获得指南）。

损伤水平评分的标准见表 11-1-11，记录找词困难子分数（范围=0～5）。

表 11-1-11 找词困难损伤水平评分标准

得分	损伤水平	依据
0	无	没有证据显示在自发性说话时出现找词困难
1	很轻微	1 或 2 次，无临床意义
2	轻微	明显的迂回陈述或同义词替换
3	中度	偶尔出现的没有进行补救的词语缺失
4	中重度	频繁的没有进行补救的词语缺失
5	重度	几乎完全丢失词语；听起来内容空洞；说出 1～2 个词

11. 理解力 这个项目用来评估受试者对口语的理解能力。评估此项时，测试者应该考虑到在测试开始时的对话和测试期间，受试者对测试者所说的话的理解程度。不包括对命令的反应。

损伤水平评分的标准见表 11-1-12，记录理解力子分数（范围=0～5）。

表 11-1-12 理解力损伤水平评分标准

得分	损伤水平	依据
0	无	没有证据显示理解力差
1	很轻微	1～2 次理解错误
2	轻微	3～5 次理解错误
3	中度	需要几次重复和改变措辞
4	中重度	受试者偶尔才能回答正确，即对于是/否的问题
5	重度	受试者很少能恰当地回答问题，这并非由于语言贫乏

12. 数字划销测验

(1) 指导语：将图 11-1-2 中的练习题放在受试者面前。说：“这张纸上面是 2 个数字，下面一串数字混合了上面 2 个数字，你从这里开始（测试者手

“6”和“1”

1 2 2 4 5 9 5 6 6 9 1 9 6 7 8 3 2 4 3 7 2 1 4 2 2 1 2 6 6 3

图 11-1-2 数字划销测验练习题

指第一行开始的数字),划掉与上面2个数字一样的数字。愈快愈好。"30 s后中止练习题。

然后将正式测验题(图11-1-3)放在受试者面前,说:"这张纸上面是2个数字,下面的一系列数字混合了上面这2个数字,你从这里开始(测试者手指第一行开始的数字),一个一个地看,只要与上面2个数字一样的数字就划掉。愈快愈好。"30 s后中止练习题。假如受试者划去第一个数字就不正确,可以说:"这是正确的要划销的数字",手指划销的目标数字。假如在测验过程中,受试者弄错了要划去的数字或停下来了,测试者应该重复上述指导语。45 s后中止测验。

"3"和"7"

1 9 8 1 3 2 3 9 4 2 9 3 2 6 4 4 6 6 1 3 8 8 4 2 7 1 4 2 8 5?

8 1 9 6 9 7 1 4 5 5 6 7 6 6 5 5 7 2 2 2 5 9 5 4 7 1 7 6 7 5?

4 9 8 8 1 4 3 4 4 6 8 4 1 4 1 7 2 4 7 9 2 6 1 7 2 1 3 4 2 3?

2 6 4 1 3 7 1 5 6 8 5 4 9 8 2 5 2 2 1 6 9 3 1 4 8 7 9 3 4 1?

2 6 1 6 9 4 2 4 2 3 2 6 3 1 4 2 8 6 9 1 7 1 3 4 3 9 8 6 5 1?

8 3 4 2 6 9 9 6 1 6 9 3 8 7 2 5 4 4 8 8 6 4 1 4 7 2 6 8 7 5?

6 3 2 6 4 4 6 8 4 5 4 6 8 9 7 3 6 8 6 5 4 9 2 5 3 5 4 7 3 5?

4 9 5 3 8 1 8 2 2 6 2 6 6 1 5 2 4 2 9 4 9 7 6 1 5 1 4 1 8 8

图11-1-3　数字划销测验正式题

(2) 评分指标:① 划销目标数字的个数;② 错误个数;③ 测验中提醒次数。

(二) 非认知行为

评定下列项目时,需查看患者晤谈期间及其前一周的情况。

1. 恐惧　询问患者或知情者恐惧发生的频度,并进行评分。评分标准见表11-1-13。

表11-1-13　恐惧损伤水平评分标准

得分	损伤水平	依　　据
1	很轻	1周1次或仅在测验时发生
2	轻度	1周内发生2~3次

（续表）

得分	损伤水平	依据
3	中度	1周内发生次数介于3、4分之间
4	中重度	几乎每天要频繁地喊出声来
5	重度	每天频繁地喊出声，每次喊叫持续时间延长

2. 抑郁　询问患者或知情者是否感到悲哀、沮丧、失望。假如回答“是”，则进一步询问抑郁的严重度与弥漫性，如活动的兴趣是否丧失、对周围环境发生的事件是否无动于衷。测试者还要观察记录患者的面容与对鼓励及玩笑的反应情况。评分标准见表11－1－14。

表11－1－14　抑郁的损伤水平评分标准

得分	损伤水平	依据
1	有轻微恶劣心境	临床上可以观察到
2	轻度	表现和报告轻度恶劣心境，兴趣部分丧失
3	中度	时常感到中度的恶劣心境
4	中重度	几乎一直有恶劣心境，相当程度的活动减少，兴趣丧失
5	重度	恶劣心境弥漫且严重，完全无动于衷，兴趣全然丧失

3. 集中注意力或分心　评定患者对无关刺激和（或）由于缺乏训练而需要在测验时重新领会指导语时表现出来的分心情况，也就是患者驾驭自己思路的能力。评分标准见表11－1－15。

表11－1－15　集中注意力或分心的损伤水平评分标准

得分	损伤水平	依据
1	很轻	只有1次注意力不集中
2	轻度	注意力不集中或分心有2～3次
3	中度	注意力不集中或分心程度介于3、4分之间
4	中重度	整个晤谈过程大部分时间注意力不集中和（或）分心频繁出现
5	重度	集中注意力极度困难

4. 对测验不合作　该项目用于评定患者对晤谈某些方面的敌对程度。评分标准见表11－1－16。

5. 妄想　包括妄想新年、先占观念及对患者行为的影响。评分标准见表11－1－17。

6. 幻觉　包括幻听、幻视、幻触，对幻觉的频度与危害性做出评估，评分标准见表11－1－18。

表 11-1-16　对测验不合作的损伤水平评分标准

得分	损伤水平	依　　据
1	很轻	仅有 1 次不合作
2	轻度	有 2~3 次不合作
3	中度	不合作次数介于 2、4 分之间
4	中重度	需要频繁的引导才能完成晤谈
5	重度	拒绝继续晤谈

表 11-1-17　妄想的损伤水平评分标准

得分	损伤水平	依　　据
1	很轻	一过性妄想信念
2	轻度	存在肯定的妄想但患者对此持有疑虑
3	中度	患者相信妄想但不影响行为
4	中重度	妄想影响了行为
5	重度	基于妄想的行为非常明显

表 11-1-18　幻觉的损伤水平评分标准

得分	损伤水平	依　　据
1	很轻	听到说一个单词的声音，一次性幻视
2	轻度	幻觉出现的时间介于 1、3 分之间
3	中度	幻觉占据了一天中的许多时间，并干扰患者的正常功能
4	中重度	幻觉出现的时间介于 3、5 分之间
5	重度	幻觉几乎持续存在，完全破坏了正常功能

7. 步态　必须对正常躯体活动与过分的来回走动做出区分。评分标准间表 11-1-19。

表 11-1-19　步态的损伤水平评分标准

得分	损伤水平	依　　据
1	很轻	极少发生
2	轻度	每天走动的次数介于 1、3 分之间
3	中度	每天频繁走动
4	中重度	每天相当频繁地走动
5	重度	不能静坐，必须过分走动

8. 运动增加　与患者的正常活动水平或以往的基础水平比较后做出评定。评定标准见表 11-1-20。

表 11-1-20 运动增加的损伤水平评分标准

得分	损伤水平	依据
1	极轻	运动量极轻微的增加
2	轻度	运动量轻微的增加
3	中度	运动量显著增加
4	中重度	运动量极显著增加
5	重度	患者必须持续运动,极少静坐

9. *震颤* 测试时让患者平伸双手、张开手指,维持约 10 s。评分标准见表 11-1-21。

表 11-1-21 运动增加的损伤水平评分标准

得分	损伤水平	依据
1	极轻	极轻微的震颤,刚好可以观察到
2	轻度	可观察到的震颤
3	中度	可观察到明显的震颤
4	中重度	可观察到非常明显的震颤
5	重度	以相当大的移位快速运动

10. *食欲增加或减少* 包括与抑郁相关的食欲变化,应与患者平时的饮食量比较。评分标准见表 11-1-22。

表 11-1-22 食欲减少或增加的损伤水平评分标准

得分	损伤水平	依据
1	极轻	轻微改变,临床可疑
2	轻度	值得注意的改变,患者不需鼓励仍能进食
3	中度	明显改变,需要鼓励才能进食;患者要求更多的食品
4	中重度	显著改变,需要不断鼓励才能进食;患者要求大量的食品
5	重度	患者不进食,需要强制喂饲;患者不断抱怨饥饿,尽管用了足够的饭量

第二节 临床痴呆评定量表(CDR)

临床痴呆评定量表(clinical dementia rating scale, CDR)包括记忆、定向、判断和解决问题、工作及社交能力、家庭生活和爱好、独立生活能力 6 个项目,做出"正常 CDR=0、可疑痴呆CDR=0.5、轻度痴呆 CDR=1、中度痴呆

CDR=2、重度痴呆 CDR=3”这 5 级判断(表 11-2-1)。有的研究者单纯以 CDR=0.5 直接作为 MCI 的诊断标准,这是否合理可行?新加坡李等(2006)以 CDR=0.5 作为 MCI 的诊断标准,采用综合一项自我报告记忆问题、4 项复杂工具性日常生活能力量表作为简短的筛查工具,识别 MCI 的准确性 87.2%,高于 MMSE 的 67.6%,与逻辑记忆延迟回忆的 86.1%非常接近。Saxton (2009)调查 3 063 例社区非痴呆老人,每 6 个月随访一次,共随访 6 年,使用 CDR=0.5 和标准化成套神经心理测验这两种不同的诊断标准,40.2%符合 CDR 标准、28.2%符合测验标准(16.6%为 aMCI、11.6%为 naMCI),15.7%被两种诊断标准共同确认为 MCI、47.4%共同确认为正常老人。认知检测在正常范围而被 CDR 判断为 MCI 的老人一般教育程度较低、测验得分较低、更多工具性日常生活能力(IADLs)损害、更多抑郁症状和更多健康主诉。测验确认的 MCI 比 CDR 有更高的 AD 转化率。笔者认为采用 CDR=0.5 诊断 MCI 有太高的假阳性而不适合作为 MCI 的诊断标准。

表 11-2-1 临床痴呆评定量表 (CDR)

能力	受损程度				
	没有(0)	可疑(0.5)	轻度(1)	中度(2)	严重(3)
记忆力	记忆力没有减退,或轻微性质的偶尔健忘	经常性的轻度健忘,事情只能部分想起:“良性”健忘	中度记忆力减退;对最近的事尤其不容易记起;会影响日常生活	严重记忆力减退;只有高度重复学过的事物才会记得;新学的东西都很快会忘记	记忆力严重减退,只能记得片段
定向力	定向力能力完好	完全能定向,但涉及时间关联性时,稍有困难	涉及时间关联性时,有中度困难。检查时对地点仍有定向能力;但在某些场合可能有地理定向能力障碍	涉及时间关联性时,有严重困难;通常对时间无定向能力,经常对地点无定向能力	只有对患者自己的定向能力
判断与解决问题的能力	日常问题、财务及商业性事务都能解决得很好;和以前的表现相比较判断力良好	在解决问题、分析事物的类似性及差异性时能力稍有减弱	在解决问题、分析事物的类似性及差异性时有中度困难;通常还能维持社交事务判断力	在解决问题、分析事物的类似性及差异性时有严重障碍;社交判断能力通常已受影响	不能做判断或解决问题

（续表）

能力	受损程度				
	没有(0)	可疑(0.5)	轻度(1)	中度(2)	严重(3)
社区事务	和平常一样能独立处理工作、购物、义务劳动及社会群体活动	参加这些活动稍有障碍	虽然还能参与这些活动，但无法独立参与；初看起来似乎还正常	在家庭以外无独立行事能力的可能性。外表足够正常，能够带去家庭以外的场所参加社交活动	在家庭以外无独立行事能力的可能性。外表看上去病情严重，不能够带去家庭以外的场所参加社交活动
家务与业余爱好	家居生活、业余爱好、知识兴趣都维持良好	家居生活、业余爱好、知识兴趣方面稍有障碍	居家生活能力方面已明显出现轻度障碍；已经放弃做较困难的家务；已放弃比较复杂的业余爱好及兴趣	只有简单家务还能做；兴趣很少，也维持得不好	在家已基本没有能力做事情了
生活自理	完全能自理		须旁人督促或提醒	穿衣、个人卫生及个人事务之料理，都需要帮忙	个人自理方面依赖别人给予很大帮助；经常大小便失禁

注：评定分数只表示与从前正常水平相比认知能力的下降，而非其他因素所造成的损害。

CDR使用说明：CDR不能用于躯体残疾或抑郁患者；6个条目中，记忆项目是主要的，另5项是次要的；当3项或3项以上与记忆项相同时，CDR＝记忆项得分；当3项或3项以上在记忆项一侧时，CDR＝项目得分相同最多的分数；当记忆项两侧分布分别为2项或3项时，CDR＝记忆项得分。

然而，除了CDR-GS(CDR global score)＝0.5这种分析方法外，目前国际上更流行的是CDR-SB(clinical dementia rating scale sum of boxes)得分指标。CDR-SB总分是将6个项目的得分简单相加之和。O'Bryant(2010)比较5 115例正常老人、2 551例MCI患者与4 796例痴呆患者，发现CDR-SB能够有效地区分这3个组别。CDR-SB＝0表示受试者正常，0.5～4.0分为可疑认知受损(其中0.5～2.0分为可疑受损，2.5～4.0分为极轻痴呆)，4.5～9.0分为轻度痴呆，9.5～15.5分为中度痴呆，16.0～18.0分为重度痴呆。在临床药物试验的疗效评估上，CDR-GS与CDR-SB都是重要的评估指标。

根据额颞叶退行性变修订的CDR(FTLD-modified clinical dementia rating scale)是在原来CDR基础上增加了“行为紊乱”和“语言评估”这两项，已经证实可以有效反映FTLD的病情严重度。

对于晚期痴呆，还可将CDR延伸至完全痴呆(CDR＝4)和终末期痴呆(CDR＝5)。完全痴呆(profoundly demented，CDR＝4)：发音或语言理解严

重受损，需辅助行走，自主进食、认识家人或控制二便困难；终末痴呆（terminal demented，CDR＝5）：无法交流、卧床、植物状态、二便失禁。

由于完成 CDR（表 11－2－2）耗时长，需要约 30 min，Duara 于 2010 年发表了 CDR 的修订版（mCDR）（表 11－2－3），其全部项目是测试者根据知情者回答获得的信息进行评分，耗时约 10 min。注意 mCDR 的部分项目与 CDR 的检测项目不一致。

表 11－2－2　临床痴呆评定问卷检查内容

此问卷提供一组基本问题来引导面谈，请提问下列所有问题。如有需要，可多问一些问题来评定患者的临床痴呆评定（CDR）情况。请将通过附加问题所获取的有关信息记录下来。

一、由知情者回答的有关记忆力的问题

1. 他/她的记忆力或思维存在问题吗？　□是　□否

1a. 如果是，该问题是不是经常出现？　□是　□否

2. 他/她能回忆起最近发生的事件（或有个人意义的事情）吗？　□通常　□有时　□很少

3. 他/她能记住简短的准备购物的清单吗？　□通常　□有时　□很少

4. 在过去的一年内他/她的记忆力有无减退？　□是　□否

5. 他/她记忆力的受损情况有没有严重到这样一种程度，以至于他/她现在完成日常活动（或退休前的活动）的能力不如几年以前了？（被访者间接看法）　□是　□否

6. 他/她是否在几周内会完全忘记刚刚发生的重大事件（如：旅行、聚会、家庭成员婚礼）呢？　□通常　□有时　□很少

7. 他/她会忘记与该重大事件有关的细节吗？　□通常　□有时　□很少

8. 他/她会完全忘记很早以前的重要事情吗？（如：生日、结婚纪念日、就业地点）？　□通常　□有时　□很少

9. 请告诉我最近在他/她生活当中发生的一些他/她应该记得的事件或活动（目的是为获取该事情发生的地点、时间、参加者、历时、完结时间和患者或其他参加者怎样到达场地等详细资料作为稍后的测试之用）。

一周内：＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

一个月内：＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

10. 他/她的出生日期？＿＿＿＿＿＿＿＿＿＿

11. 他/她的出生地？＿＿＿＿＿＿＿＿＿＿

12. 他/她上的最后一所学校是哪一所？＿＿＿＿＿＿＿＿＿＿

校名＿＿＿＿＿＿＿＿＿＿

地点＿＿＿＿＿＿＿＿＿＿

年级＿＿＿＿＿＿＿＿＿＿

13. 他/她以前的主要职业/工作是什么（如果患者没有职业，他/她配偶以前的主要工作是什么）？

＿＿＿＿＿＿

14. 他/她最后一个工作单位是什么（如果患者没有职业，他/她配偶的最后一个工作单位是什么）？

＿＿＿＿＿＿

15. 他/她（或配偶）什么时候退休的？为什么退休？

二、由知情者回答的有关定向力的问题

他/她是否能够准确知道：

1. 今天是几号？

□经常　□有时　□很少　□我不知道

（续表）

2. 月份？

□经常 □有时 □很少 □我不知道

3. 年份？

□经常 □有时 □很少 □我不知道

4. 星期几？

□经常 □有时 □很少 □我不知道

5. 他/她在判断事件的时间顺序上存在困难吗（当事情按时间顺序发生时）？

□经常 □有时 □很少 □我不知道

6. 他/她在熟悉的街道能找得着路吗？

□经常 □有时 □很少 □我不知道

7. 在自己家居住地区以外的地方，他/她是否能知道怎么从一个地方到另一个地方？

□经常 □有时 □很少 □我不知道

8. 他/她在熟悉的建筑物内能找到路吗？

□经常 □有时 □很少 □我不知道

三、由知情者回答的有关判断和解决问题能力的问题

1. 总体来说，如果你要评价他/她现时的解决问题的能力，你认为他/她的能力：

□和以前一样好

□好，但是不如以前

□一般

□差

□完全丧失

2. 评估他/她处理少量金钱（如：找钱、根据重量计算商品价钱）的能力：

□没有丧失

□部分丧失

□严重丧失

3. 评估他/她在处理家庭钱财方面的能力（如：到银行存款、取款，支付水电费等）：

□没有丧失

□部分丧失

□严重丧失

4. 他/她能够有效处理家中的急事（如：水管漏水或发生小火灾）吗？

□与从前一样好

□因为思维问题而比以前差

□因为其他原因而不如以前，这些原因是：________________

5. 他/她能够理解不同的场合或他人的解释吗？

□经常 □有时 □很少 □我不知道

6. 他/她在社交场合或与人交往中能够举止正常吗（即，与他/她发病前的举止比较）（此项目是评估行为，而不是外表）？

□经常 □有时 □很少 □我不知道

四、由知情者回答的有关社区事务能力的问题

职业方面

1. 患者是否仍然在工作？ □是 □否 □不适用

如不适用，跳到问题 4

如果是，跳到问题 3

如果不是，跳到问题 2

2. 记忆力或思维问题是他/她决定退休的原因之一吗？（跳到问题 4）

□是 □否 □我不知道

3. 患者是否由于记忆力或思维问题而在工作上出现重大的困难？

□很少或从不 □有时 □通常 □我不知道

（续表）

社交方面

4. 他/她以前骑过自行车吗？ □是 □否

他/她现在还骑自行车吗？ □是 □否

假若没有，是否是因为记忆力或思维问题？ □是 □否

5. 如果他/她仍坚持骑自行车，他/她在骑自行车时会因为思维能力不佳而出现什么问题或危险吗？ □是 □否

6. 他/她能独立地根据需要去购物吗（如果需要表明患者在该方面的能力水平，请注明）？

□很少或从不　□有时　□通常　□我不知道

（购买任何东西时都需要人陪伴）　（购物数量有限，会买重或者忘记需要购买的物品）

7. 他/她能独立地从事家庭以外的活动吗？

□很少或从不　□有时　□通常　□我不知道

（在没有人协助的情况一般不能进行活动）（有限的活动和/或例行活动例如象征性参加社交活动或会议，到发廊理发等）　（有目的地参与一些活动，例如选举投票）

8. 他/她有没有被带去参加家庭以外的社交活动？ □是 □否

如没有，为什么？________

9. 与患者短暂接触的人会觉得他/她有病吗？ □是 □否

10. 如果是在疗养院里，他/她能够很好地参与社交活动吗（智力上参与）？

□是 □否

重要提示：

现在有没有足够的信息去评估患者在社区事务方面的能力衰退程度呢？如果没有，请进一步往下调查。

社区事务：例如参加社交活动、拜访朋友或亲人、参加政治活动、参加行业组织（如律师协会）、其他行业团体、社会俱乐部、服务机构、教育项目。

五、由知情者回答的有关家务和爱好能力的问题

1a. 他/她在做家务方面的能力出现了什么变化？________

____

1b. 有哪些家务他/她现在还仍然能做得好？________

____

2a. 他/她在从事业余爱好方面的能力发生了什么变化？________

____

2b. 有哪些业余爱好他/她现在仍然能做得好？________

____

3. 如果在疗养院，有哪些家务和爱好他/她再也无法做好了？________

____

日常活动

	没有丧失		严重丧失
4. 处理家务方面的能力	0	0.5	1

请描述：________

____

5. 他/她在做家务方面的水平是：

（选一个，不需要直接向知情者即提供有关患者的信息的人发问）。

□缺乏有目的地去做事的能力。

（只有经过许多指导才能做诸如铺床之类的简单工作）

□只能在有限活动中发挥一些作用。

（在一些指导下，洗碗碟能达到可接受的清洁程度，在桌子上摆放餐具）

□能够独立完成某些工作。

（操作诸如吸尘器之类的家用电器或准备便餐）

□能够从事平常的活动，但达不到平常的水平。

□在平常活动中表现正常。

（续表）

重要提示：

现在有没有足够的信息去评估患者在家务和爱好方面能力的衰退程度呢？如果没有，请进一步往下调查。

家务：做饭、洗衣服、清洁、购买食品杂货、倒垃圾、打理花草、家庭用具简单维护和基本维修。

业余爱好：缝纫、绘画、手工艺、阅读、娱乐、摄影、园艺、去看电影或欣赏音乐会、木工、参与体育运动。

六、由知情者回答的有关个人自理能力的问题

＊请你评估他/她在以下各方面的心智能力如何：

	不需要帮助	偶尔出现扣错纽扣等之类的错误	次序颠倒一般会忘记一些东西	不能穿衣
A. 穿衣	0	1	2	3
	不需要帮助	需要提示	有时需要帮助	是或几乎总是需要帮助
B. 盥洗打扮	0	1	2	3
	能干净地进食，并适当地使用餐具	只会用勺子而且用得一团糟	简单的固体食物（如馒头）	完全需要别人喂食
C. 进食	0	1	2	3
	控制完全正常	有时会尿床	频繁尿床	大小便失禁
D. 大小便控制	0	1	2	3

＊如果患者个人自理能力的水平比以前差，即使他们不需要提示，也可以考虑在此表上打1分。

七、由患者回答的有关记忆力的问题

1. 你在记忆力和思维方面有没有什么问题？　□是　□否

2. 刚才，你的（配偶等）告诉了我你近来经历的一些事件或活动。你可否告诉我一些与这些事情有关的细节？（如有需要可询问细节，如：事情发生地点、时间、参与人员、历时多久、何时结束、患者或其他的参加者是如何到达场地的）。

一周内

1.0—大部分正确 ________

0.5 ________

0.0—大部分不正确 ________

一个月内

1.0—大部分正确 ________

0.5 ________

0.0—大部分不正确 ________

3. 我会给你一个名字及地址，请你记住它们几分钟。请跟我重复：

（重复提示直至所给名字和地址被记住或最多3次）

项目	1	2	3	4	5
	张	三	上海市	南京路	42号
	张	三	上海市	南京路	42号
	张	三	上海市	南京路	42号

（在每次重复正确的部分下面画线）

4. 你的出生日期？________

5. 你的出生地？________

（续表）

6. 你就读的最后一所学校是哪一所？______
校名______
地点______ 年级______
7. 你以前的主要职业/工作是什么(如果您无业,您配偶以前的主要工作是什么)？______
8. 你最后一个工作单位是什么(如果您无业,您配偶的最后一个工作单位是什么)？______
9. 你(或配偶)何时退休的？为什么退休？______
10. 请重复刚才我让你记住的姓名和地址(在被正确重复的部分下画线)：

项目	1	2	3	4	5
	张	三	上海市	南京路	42 号

(在每次重复正确的部分下面画线)

八、由患者回答的有关定向力的问题

逐字记录患者所给每题的答案

1. 今天是几号？ □正确 □不正确
______
2. 今天是星期几？ □正确 □不正确
______
3. 现在是几月份？ □正确 □不正确
______
4. 现在是哪一年？ □正确 □不正确
______
5. 你现在所处的场所叫什么名字？ □正确 □不正确
______
6. 我们现在位于哪个城镇或城市？ □正确 □不正确
______
7. 现在时间是几点了？ □正确 □不正确
______
8. 患者是否认识知情者是谁(根据你本人的判断)？ □正确 □不正确
______

九、由患者回答的有关判断和解决问题能力的问题

说明：如果患者最初的反应达不到零分,需要进一步询问,以便鉴别患者对该问题理解的最佳程度。圈出最接近的答案。

相似：

例如："钢笔和铅笔有什么相似之处？(书写工具)
以下这些东西有什么相似之处？" 患者的回答

1. 白萝卜……椰菜花 ______
(0＝蔬菜)
(1＝能吃的食物、生物、可煮食等)
(2＝答语不相关、不同、购买它们)
2. 书桌……书柜 ______
(0＝家具、办公家具；全用于放书)
(1＝木头、腿)
(2＝不相关、不同)

区别：

例如："糖和醋有什么不同？(甜与酸)
以下这些东西有什么不同之处？"

3. 撒谎……错误 ______
(0＝一个故意、一个无意)
(1＝一个不好另一个好,或只解释了一个)

（续表）

(2=任何其他答案、相似点)
4. 河……运河 ____________________
(0=天然、人工)
(2=任何其他答案)
计算：
5. 1元里有多少5分？ □正确 □不正确
6. 5.40元里有多少个2角？ □正确 □不正确
7. 用20减3，再用每一个得数接着减3，一直减下去，请给出过程和结果。
□正确 □不正确
判断：
8. 当你在没有准备的情况下到达一个陌生的城市，你如何去找到你想见到的朋友？
(0=查询电话号码簿，到当地户政部门查询，打电话给共同认识的朋友)
(1=打电话给公安局或派出所，打114查询但通常不会给地址)
(2=没有明确的答复)
9. 患者对自身能力缺陷和社会地位的评估，以及对他/她参加这次评估的原因的了解程度
□有良好认识 □有部分认识 □几乎无认识

表 11-2-3 mCDR的检查内容

1. 记忆力	① 最近事件的记忆；② 即使中断任务也能完成；③ 记人名；④ 交谈中词语的运用；⑤ 财产所有权的记忆；⑥ 义务的记忆
2. 定向力	① 事件之间的关系；② 找到邻近地区的位置；③ 找到非邻近地区的位置；④ 日期记忆；⑤ 月份记忆；⑥ 年份记忆
3. 判断与解决问题的能力	① 对于错误的自我纠正；② 适当的小心谨慎；③ 计划与组织；④ 购物的找零；⑤ 处理财务；⑥ 判断所购之物是否合适
4. 社区事务	① 社区活动；② 个人卫生；③ 冲动；④ 兴趣缺乏；⑤ 设身处地、将心比心；⑥ 对社交情境的反应
5. 家务与业余爱好	① 家庭杂务；② 器具的使用；③ 处理复杂的日常事务；④ 处理复杂的非日常事务；⑤ 理解和跟踪书籍或电子传媒；⑥ 分享游戏与爱好
6. 生活自理	① 小便控制；② 大便控制；③ 穿衣；④ 盥洗打扮；⑤ 使用饮食器皿

注：跟他(她)成年时期的最佳状态相比，他(她)目前的表现是：0=没有变化；0.5=可疑的恶化；1=恶化但不需要协助；2=明显恶化，需要协助。

第三节 Mattis痴呆评定量表(DRS)

Mattis于1976年编制的痴呆评定量表(dementia rating scale，DRS)是一套标准化的临床精神状态检测工具，在西方国家广泛应用。在中国香港和上海等地有应用的报道。

一、DRS 的内容与修订

DRS 有 5 个因子分，项目与得分见表 11-3-1。其中：① 注意 37 分：包括数字广度、执行比较复杂的口头指令、数出随机排列的 7 的个数、读一组词语和图片匹配。② 起始与保持 37 分：包括言语流畅性、语言重复、两手交替运动和书写运动。③ 概念形成 39 分：包括词语归类和图片相似性。④ 结构 6 分：模仿画几何图形。⑤ 记忆 25 分：包括定向、词语再认、句子延迟回忆、无意义图案再认等。共有 37 道题目，费时 15～30 min，总分 144 分。

表 11-3-1 Mattis 痴呆评定量表

DRS 项目	满分	DRS 项目	满分
注意分测验		结构分测验	
数字广度	8	绘图设计	6
连续两个指令	2	概念化分测验	
单一言语指令	4	相似性	8
模仿	4	归纳思考	3
7 的个数	6	不同点	3
随机排列的 7 的个数	5	相似性-多选题	8
读一组词语	4	相同与不同	16
图片匹配	4	自发语言	1
分测验总分	37	分测验总分	39
起始/持续分测验		记忆分测验	
语言流畅性	20	句子回忆	4
服装名称	8	自发语言回忆	3
语言重复	2	定向	9
双重交替动作	3	词语再认	5
书写运动	4	图片再认	4
分测验总分	37	分测验总分	25
		总分	144

DRS 的翻译与修订：中国香港广东话版直接从美国英文版本翻译而来，经过回译，并由第三译者校正其偏差，项目修订情况是：①“从随机排列的大量数字中数出 7 的个数”代替“从随机排列的大量字母中数出 A 的个数”，因为中国老人对英文字母不熟悉。② 概念形成分测验中，识别 3 个词中哪两个同类，中英文版所用词语不同。③ 以“香港特区行政长官是谁?”代替“美国总统是谁?”

中文普通话版在中国香港广东话版的基础上，修改了两道题：① 由"我国现任总理是谁?"代替"香港特区行政长官是谁?"② 言语流畅性中"列举超市物品名称"改为"列举动物名称"，因为当时(1998 年作者修订时)内地超市还不普及，超市内摆放哪些商品老人可能不熟悉，故予修改。

DRS 普通话版的信度与效度：DRS 中文版的效度良好，与 MMSE 的相关系数为 0.71。不管是中文版还是英文版，DRS 总分与老人的年龄和教育程度显著相关，与性别无关。DRS 中文普通话版根据教育程度不同制定 DRS 总分的划界分，敏感性和特异性比较理想。

二、评分

DRS 的分析指标为总分和因子分。总分 144 分，注意 37 分，起始与保持 37 分，概念形成 39 分，结构 6 分，记忆 25 分。根据国人受教育程度不同，中文版 DRS 的划界分为：文盲组 90 分，小学组 115 分，初中及以上组 120 分，低于划界分为认知功能受损。

三、评价

由于部分老年患者认知损害非常严重，无法完成全套的韦氏智力测验和韦氏记忆测验，而 MMSE 又没有因子分，DRS 量表可弥补这两者的不足。此外，与 MMSE 比较，DRS 有许多优点，如题量较大易于获得更全面的认知功能缺损与保存状况的信息。每个因子的题目由难到易排列，能完成较难的就不再做该部分中较易的题目，以节约时间，故正常老人只要 10～15 min就可完成，在各种综合性神经心理测验中属于费时最少的测验之一。由于部分题目非常简单，受试者很少出现"地板效应"，在一组 MMSE 平均为 8 分的 AD 患者中，DRS 平均 58 分，故 DRS 常被用来判断痴呆患者认知损害的严重度，也可用于中-重度患者的纵向随访和中-晚期患者的疗效评定。因而，在临床试验中被用于评价药物疗效。其"起始与保持"和"概念形成"等项目是 MMSE 没有的，被认为对额叶和额叶-皮质下功能失调较敏感，有助于痴呆的鉴别诊断(血管性痴呆和亨廷顿病以额叶-皮质下功能失调为主)。另外，DRS 的语言项目和非语言项目在数量和得分上平分秋色，Kessler 提出语言和非语言双因子模型，这是其他综合性测验没有的，有利于分析语言背景的影响和语言的损害程度。

"记忆"与"起始/持续"2 个因子是 5 个因子分中鉴别效力最好的。但这并

不是说另 3 个因子不重要，在今后的测验中可以省略，因为不同的因子反映了受试者不同侧面的认知功能。当 AD 处于不同的严重程度时，其认知功能的保存和缺损的领域也不同，如在痴呆的开始阶段，记忆损害非常明显，而另 4 个因子影响较少，随着痴呆进入中-重度，记忆损害减慢，甚至处于平台期，注意、起始/持续和概念化 3 个因子的损害却在加速。经历长期汉字结构训练的中国老人组在空间结构能力方面明显优于美国老人组的表现，故该因子对美国老人适合，对于中国文化程度较高的老人，其难度过低，因“天花板效应”而降低敏感性。

临床上为了节约时间、快速筛选 AD 型痴呆患者，可以选择性地采用“记忆”与“起始/持续”2 个因子中的部分敏感项目。由“定向、言语流畅性、句子延迟回忆”组成的“敏感因子组合”，耗时不到 5 min，不需借助任何工具，适合在普通门诊初筛时使用。事实上，这个“敏感因子组合”与笔者编制的 MES（见第三章第四节）在内容与形式方面均有相似之处。DRS 的缺点是对临床前痴呆或 MCI 的检测敏感性和特异性与 MMSE 相似，并未改善。

（郭起浩）

第四节　计算机辅助神经心理测验

计算机与信息技术的发展一日千里，它在神经心理学的应用包括：计算机辅助的认知功能评估；移动电信认知评估如电子问卷和电子测验；虚拟现实环境的应用；人工智能或按规则推理的应用。本节只简略介绍第一种情况。

神经心理测验一个令人兴奋的发展方向是计算机辅助神经心理测验，它的基础是项目反应理论，对于正常认知功能和极轻微的损害，它是简洁的、敏感的方法，但是，它的成本比执笔测验高，需要信息科学的专家参与设计与分析，对受试者的教育水平和理解能力有一定要求，这些局限性影响了它在发展中国家的应用。

目前常用的计算机辅助神经心理测验有自动神经心理评估体系（automated neuropsychological assessment metrics，ANAM）、计算机实施的 MCI 筛查（the computer-administered neuropsychological screen for mild cognitive impairment，CANS-MCI）、认知药物研究计算机评估系统（cognitive drug research computerized assessment system，COGDRAS）、计算机辅助成

套神经心理测验（computerized neuropsychological test battery，CNTB）、Cambridge 神经心理测验自动化版（Cambridge neuropsychological test automated battery，CANTAB）和 MCI 计算机评估（computer assessment of mild cognitive impairment，CAMCI）等。各种版本的项目大部分是相同的，且大部分项目是根据纸笔版本转化修订而来。

一、内容与修订

（一）CAMCI

CAMCI 有 9 个分测验，其中 7 个分测验是标准化纸笔测验的计算机版本：星星划销测验、数字广度测验、词语记忆与再认测验、图片再认测验、Go/No-Go 测验是根据纸笔版本转化修订而来的，2 个虚拟任务：虚拟现实的商店购物路径测验（包括将要购买物品的前瞻性记忆、路径的选择和偶然记忆等）和虚拟银行的 ATM 取款任务。完成 CAMCI 大约需要 20 min。

为了适应计算机呈现，计算机版本与传统的纸笔测验略有不同。

1. 星星划销测验　反映注意力。计算机屏幕随机呈现星星、圆圈、正方形、三角形，当星星出现在屏幕上的时候要求受试者敲一下键盘。

2. 数字广度测验顺背分测验　反映注意力。每秒 1 个数字，从 3 个数字到 6 个数字，要求受试者按照原来次序回忆这些数字，在屏幕下方的显示器上呈现。

3. 词语再认测验　反映言语记忆。一次性屏幕上呈现一组 6 个单词，要求受试者记住每个单词，并告知后面需要回忆。一段时间间隔后，呈现 6 个组，每组 4 个单词（3 个干扰词、1 个目标词），认准目标词敲一下键盘。

4. 词语回忆测验　反映言语记忆。一次性呈现 5 个由 3 个字母组成的单词，告知需要回忆。共呈现 3 次。大约 10 min 时间间隔后，要求受试者回忆这些单词，在键盘上打出这些单词。

5. 图片再认测验　反映视觉记忆。呈现固定系列的图片，有些是以前呈现过的，有些是新的，假如图片是呈现过的，受试者说“是”，没有呈现过的，说“不”。

6. Go/No-Go 测验　反映执行功能。第一部分，要求患者听到 1 次蜂鸣器的声音，跟着敲 2 次；听到 2 次声音，患者跟着敲 1 次。第二部分，规则改变，听到蜂鸣器叫 1 次，患者敲 2 次，而听到叫 2 次时患者不敲。

7. 数字广度测验倒背分测验　反映工作记忆。

8. 虚拟现实的商店购物路径测验　通过导航穿过平板电脑的虚拟世界。受试者驾车来到苏利文市场购买商品，在路上，不能出偏差，比如，商店位于银行与邮政局之间。测验取向有更好的生态学效度，更适应老年人的生活。相比实验室的静态评估，这种方法有机会评估受试者的日常功能，如前瞻性记忆（记住将来的事情）、偶然记忆（记住没有刻意去记的材料）和决策能力。

9. 虚拟银行的 ATM 取款任务　当受试者来到虚拟银行的标准化的 ATM 前，要求受试者取款 250 元。交易业务的每一步都进行评分，并记录完成任务的时间。假如受试者不记得取款任务，测验会自动“载”受试者到银行的 ATM 前，所以，所有受试者都会完成这部分测试。

（二）CANTAB

CANTAB 反映 3 个认知区域：工作记忆和计划、注意力和视空间记忆。通过点击触屏回答问题，并且大部分测验不需要语言指导。虽然许多报道是基于 13 个分测验的小测验，但是这套测验是目前同类测验中被使用最多的测验。基于 770 多个正常人作为对象的主成分分析确定了两个因子：广义学习和记忆、反应速度。最早报道的文章之一显示，Sahakian 和 Owen 确定了对健康对照、AD 早期患者和 PD 患者尤其灵敏的分测验配对联想学习法、延迟样本匹配任务和场景转换注意力。Fray 等早期的综述支持应用 CANTAB 评估其他神经退行性疾病。后续的研究报道了 CANTAB 的重测信度、基于健康老年人的大样本常规数据和用于记忆障碍的早期检测。

1. 图片识别（pattern recognition）　测验分为两个阶段。第一阶段在屏幕中心白色的方框里展示给受试者 12 张有颜色的图片，一次出现 1 张，每张显示 3 s。第二阶段，12 对有颜色的图片连续出现，一次出现 1 对，每对中有 1 张照片在第一阶段出现过，而另 1 张是新出现的。在识别阶段，目标图片出现的顺序和第一阶段目标图片出现的顺序相反，错误图片和目标图片在形式上有差别，但是颜色没有差别。受试者通过点击图片来选择他们认为正确的目标图片。

2. 空间识别（spatial recognition）　第一阶段，依次出现 5 个 1 英寸的白色空正方形，每次出现 1 个，且每次出现在屏幕的不同位置，每个出现 3 s。第二阶段，同时出现 2 个正方形。其中一个目标正方形出现在第一阶段出现过的位置，另一个错误正方形出现在第一阶段未出现过的位置。受试者必须识别出目标形状并去点击它。第二阶段，目标正方形出现的顺序和第一阶段目标正方形出现的顺序相反。

3. 同时延迟样本匹配任务(simultaneous delayed matching to sample) 每个测验开始的时候,会在屏幕中心出现一个复杂的抽象图形,由 4 个象限构成,每个象限的颜色和形式都不一样,每个图形出现 4.5 s。告诉受试者去研究这个图形,因为稍后会要求受试者在 3 个错误图形中识别这个图形。4 个图形同时出现在样品图形下边。要求受试者点击颜色、形状,准确匹配样品图形的图形,每次只有一个图形和样品图形匹配。其他的选择项图形都是新出现的错误图形,和样品图形在颜色和形状上都有差别。和样品图形相比,一个错误图形是颜色相同、形状不同,另一个错误图形是形状相同、颜色不同。

延迟条件和同时条件相同,样品图形 4.5 s 后从屏幕上立即消失。0、4、12 s之后 4 个选择图形出现,受试者需要做出选择。3 次练习后(同时,0 s 和 12 s),总共有 10 次测验,每次都有 4 个同时和延迟条件依次出现,并且错误图形以随机顺序出现(总共 40 次)。

4. 视空间匹配相关学习(visuospatial paired associated learning) 在这个测试中,要求受试者记住 8 个图形-位置的相关性。最初,在屏幕周围展示 6 个固体白色盒子,告诉受试者这些盒子会依次打开展示盒子里边的东西,受试者的任务就是寻找盒子里有颜色的图形并且记住每个图形属于哪个盒子。每个盒子打开 3 s 后关上,然后另一个盒子打开,以随机顺序依次打开所有的盒子。第一次测验只有一个盒子含有带颜色的图形。最后一个盒子打开后,这个图形立即出现在屏幕中心,受试者需要点击图形所在的盒子。虽然每次回答后不会立即提供反馈信息,但是如果所有选项都回答正确,"全部正确"字样会出现在屏幕中心,受试者继续进行接下来的测验。如果选项不是全部正确,盒子会重新打开 2 s 后关闭(提醒阶段),然后给予受试者第二次机会尝试去正确分配图形。每次测验,受试者有 9 次提醒的机会,总共有 10 次尝试回答的机会。如果受试者全部回答错误,这个测试就停止了,经历的情况评分就会出现在屏幕下方。如果 9 次提醒中有回答正确的,测验继续进行。一个图形的初始阶段之后,下一阶段会有一个图形,然后 2 阶段,每阶段都有 2 个图形;然后 3 阶段,每阶段都有 3 个图形;然后是 6 个图形的阶段(例如:每个盒子中有 1 个图形)。最后,两个额外的盒子出现在屏幕上,要求受试者总共正确定位 8 个图形。受试者正确回答所有的图形后自动从一个阶段进入下一阶段,或者在最初的展示阶段后,或者在任何一次提醒阶段。

5. 空间工作记忆(spatial working memory) 受试者需要在屏幕上出现的大量盒子中寻找里边隐藏有"蓝色记号"的盒子。展示的盒子中有一个盒子

里隐藏有记号，受试者需要通过依次点击、打开盒子寻找。如果含有蓝色记号的盒子被发现，那么这个盒子就不再用来藏记号。前一次找到的含有蓝色记号的盒子如果再次被打开，那么作为一次错误记录。12 次任务(每次 4 个，含有 4、6、8 个盒子)中出现的错误次数之和作为表现指标用来分析。

6. 匹配样本(视觉寻找指定目标)(matching to sample-visual search) 屏幕上有 8 个白色盒子围绕着中心的 1 个红色盒子。每次测试开始的时候，要求受试者把手放在键盘上。一旦受试者按下键，盒子就被打开显露中心的目标刺激，它被选择刺激围绕，选择刺激中有一个和目标刺激是一样的，把它找出来。刚开始的时候，告诉受试者这些任务，并且告诉受试者尽可能快地松下按键，然后在周边的盒子中点击和中心样品一样的目标刺激。屏幕上会显示正确和错误的信息。在试验中分别有 1、2、4、8 个不同的样本可被选择，比例是相同的。开始 48 个测试前，有 4 个样本量逐渐增大的例子。测试刺激和延迟匹配样本测试中的类似。每个测试刺激都有 4 个按照颜色和图形变化的象限组成，但是在给定的选择集中有相同的 4 个颜色。当数据集大于 1 时，通过变化象限的相对位置，刺激一半来自目标刺激，一半来自错误刺激。

二、评价

CAMCI 是一种便携式的基于电脑的测验，包括一项革命性的虚拟环境任务以及根据电脑改良的神经心理测验。它可以有效评估老年人早期认知功能障碍。一项针对 524 名非痴呆的社区居民的研究显示，CAMCI 鉴别轻度认知功能障碍的灵敏度高达 86%，特异度高达 94%。

设计 CAMCI 是为了解决临床上认知功能测验的问题，协助评估怀疑有认知功能障碍的成年患者以及对这些患者的进一步评估，但是 CAMCI 不提供医疗诊断。CAMCI 是一项受试者可以自己进行评估的测验，虽然需要在医疗专业人员的指导下进行，但是 CAMCI 基于电脑，由患者自主控制速度(即按照自己的速度一项一项进行测验；即使受试者不会使用电脑也能完成 CAMCI 测验)，自动评分，在训练有素的专业人员初始设置之后不需要专业人员一直在场。受试者完成测验的平均时间为 25 min。受试者的报告可以立刻在屏幕上生成以供查看，或者打印报告供医生、合格的卫生保健专业人员进行分析。受试者的表现根据常模，按照年龄和教育进行调整，最终得到一个综合分数及各分测验的分数，同时给出任务完成精度和反应时间的分数。测试在平板电脑上操作，比台式电脑有更大的便携性和灵活性。这种方法可以确保

标准化管理和评分，从而避免单位之间和检查者之间的变异性。

CAMCI 已经被用于下列研究：初级保健患者的认知评估研究、痴呆干预、创伤性脑损伤、与艾滋病相关的认知能力下降、运动评估、运动干预，以及老年人和关节炎的研究。

CAMCI 优点：高敏感度和高特异度，基于计算机，患者自我控制速度，准确度和反应时间的跟踪和报道，可立即在屏幕上查看报告或打印报告，容易使用，报告包括综合得分和各个任务的分数，在医疗级平板电脑预装，便携，手写笔输入，有集成扬声器。

CAMCI 缺点：不能用于无法矫正的视觉或听觉障碍患者，肢体障碍以至于无法使用触屏者。CAMCI 不适用于居家使用。

（梁小妞　郭起浩）

第五节　严重损害量表(SIB)

严重损害量表(severe impairment battery, SIB)由 Saxton 等于 1990 年编制，量表包含 51 个项目，包括定向力、注意力、语言、运用能力、视知觉、记忆力、空间结构、呼名回应和社会交往等 9 个因子，同时还包括详细的行为评估，耗时约 30 min，总分范围为 0～100(表 11-5-1)。评分愈低，说明痴呆程度越严重。SIB 重测信度 0.8，测验者之间信度 0.99。此量表目前是评价中重度到重度阿尔茨海默病(AD)疗效的最常用量表，已应用于美金刚的临床试验，也作为多奈哌齐或卡巴拉汀治疗中重度 AD 的治疗效果评估工具。

表 11-5-1　严重损害量表(SIB)

社会交际

1. (SI)

a) 接近受试者并做出要和对方握手的表示，同时口中说："您好，我叫……"

☐ 2　自发与测试者握手

☐ 1　起立，有与测试者握手的倾向，但未接触到测试者的手

b) 向一间办公室或桌子做手势并伸出一只手臂，同时说："我希望您回答我一些问题"，再说："跟我(到办公室里)来或到这边来。"

如果受试者没有反应，可以搀扶受试者的手臂，再说："跟我来。"

如果受试者不能行走，说："我希望您回答我一些问题，您能坐下/回去/过来吗?"

如果受试者没有反应，可以搀扶受试者的手臂，并说："请坐下/回去/过来。"

☐ 2　按照指令自动向相应的方向移动或自动地坐下/回去/过去

☐ 1　在测试者以搀扶示意后才做动作

（续表）

c）伸出手臂并指示一张椅子，同时说："请坐这儿。"

如果没有反应，可搀扶受试者的手臂并指示其坐在椅子上，说："坐这儿。"

如果患者当时坐在轮椅内，还可以说："请到这张桌子旁边坐。"

如果没有反应，可以将你的手温柔地放在受试者的肩膀上，并说："请把桌子拉到您的旁边。"

如果还没反应，可以再用手拍拍那张桌子，并重复上述指令。

☐ 2　自动坐在椅子上或者自动将自己的轮椅转到桌子旁，或者自动把桌子推到椅子旁

☐ 1　在测试者以搀扶示意后才行动

记忆力

2.（M）说："我叫……"（只说名或姓，可包括身份称谓，如卡尔或史密斯、先生或太太）

重复名字说："我希望您能记住我的名字，因为我待会儿还要问您"，（暂停）再说"我叫什么名字？"然后，不论答案正确与否都说："是的，我的名字是________。"

☐ 2　自动说出正确答案

☐ 1　所说答案比较接近正确答案（如以朱莉代替朱蒂）

定向力

3.（O）说："您叫什么名字？"

如果受试者只说出自己的姓或名，则再问其未说出的部分，如"约翰什么？"

☐ 2　说出全名，其间可提醒一次

☐ 1　只能说出姓或名，或原用名

语言能力

4.（L）

a）说："请在这里写下您的名字。"

☐ 2　自动写下正确的名字（可以允许签名中存在某些简写甚至缩写，特别是当受试者按照其平时的习惯来签名时）

☐ 1　部分正确，如签名中只有姓或名，或者为受试者的原用名

b）如果受试者在回答 4a 题时已得到 2 分，则跳过此题，并给予满分（2 分）。

在黑色的纸上打印受试者的姓名，并说："您能将这些抄写下来吗？"

☐ 2　自动正确抄写（打印体姓名或签名）或 4a 题回答正确

☐ 1　部分正确

定向力

5.（O）说："现在是几月份？"

如果受试者无反应，则给予提示说："现在是__月、__月还是__月呢？"

所给的备选月份应分别是 6 个月前、当前和下个月的月份

☐ 2　自动说出正确答案

☐ 1　在给出多选提示后才说出正确答案

语言能力

6.（L）说："告诉我一年中有哪几个月？"

如果受试者没有反应，则提示说："一年以一月、二月和三月开始，然后是__月？"

☐ 2　自动说出正确答案

☐ 1　在提示后说出正确答案，或者仅漏掉 1 或 2 个月份（可以给受试者 2 次提示）

定向力

7.（O）说："这座城市叫什么名字？"

如果受试者没有反应，则提示说："这是________，________，或________（城市名）吗？"提示时给出正确答案的城市名和两个其他的城市名作为备选答案。

☐ 2　自动说出正确答案

☐ 1　在给出多选提示后说出正确答案

语言能力

8.（L）

a）说："您如何称呼您平时用来喝咖啡的东西？"

如果受试者没有反应，则提示说："您用来喝咖啡的瓷器/物件/陶器叫什么？"

（续表）

☐ 2 答“杯子”或“茶杯”
☐ 1 说出某些与正确答案相关的词汇，如“玻璃杯”或“咖啡壶”，或在提示下说出正确答案
☐ 0 说出某些与正确答案不相关的词汇，如“盘子”

b) 说：“您如何称呼平时你用来盛汤的东西？”

如果受试者没有反应，则提示说：“您用来喝汤的银质物品/铜质物品/器具叫什么？”

☐ 2 勺子
☐ 1 说出某些与正确答案相关的词汇，如“汤碗”，或在提示下说出正确答案
☐ 0 说出某些与正确答案不相关的词汇，如“小刀”

9. (L)

a) 向患者呈现写有“把您的手给我”的卡片，确保患者的注意力已集中于这张卡片上，说：“请阅读这张卡片上的字并按照文字的要求做相应的动作。”

如果受试者没有反应，则通过重复上述指令的方法给予提示，同时向受试者伸出测试者自己的手，张开手掌。

如果受试者仍无反应，则大声阅读卡片上的内容。

☐ 2 受试者自动给出自己的手
☐ 1 受试者做出较接近题目要求的动作，如抬高自己的手等；或者在提示后做出正确的动作
☐ 0 当测试者不得不自己阅读卡片上的内容时

b) 说：“现在给我您的另一只手。”

如果受试者没有反应，可重复上述指令，由测试者做手势张开自己的手。

☐ 2 受试者自动给出自己的另一只手
☐ 1 受试者做出较接近题目要求的动作，如抬高自己的手但是却没有将手移向测试者；或者仍将与上题中相同的手交给测试者；或者在提示后做出正确的动作

c) 再次向受试者呈现写有“把您的手给我”的卡片，并说：“这上面说的是什么？”

如果受试者没有反应，可提示说：“大声念出这张卡片上的内容”，再拿走卡片

☐ 2 自动阅读卡片上的内容
☐ 1 部分正确，如读错了卡片上的内容或者只读出卡片中句子的一部分，或者在提示后做出正确的反应

记忆力

10. (M) 说：“对不起，刚才您说的是什么？”

如果受试者没有反应，可提示说：“你说了什么？”

☐ 2 受试者自动正确地重复在自己 9c 中说过的话。
☐ 1 部分正确地重复自己先前的话。即只重复句子的一部分或在提示后正确重复刚才的话

语言能力

11. (L) 说：“现在说这个”

a) 说：“人们花钱。”

☐ 2 正确重复
☐ 1 部分正确地重复，或者用该词汇说出评论性的语句，如“钱永远是不够的。”

b) 说：“婴儿。”

☐ 2 正确重复
☐ 1 部分正确地重复，或者用该词汇说出评论性的语句，如“我喜爱婴儿。”

注意力

12. (ATT) 说：“现在说数字。”

“2”	“5”	“87”	“41”		
☐	☐	☐	☐		
“582”	“694”	“6439”	“7286”	“42731”	“75836”
☐	☐	☐	☐	☐	☐

（续表）

如果受试者没能正确地重复两个相同位数的数字，则停止此项测试。

□ 2　正确重复含有 3 个、4 个或 5 个数字的数字串

□ 1　正确重复含有 1 个或 2 个数字的数字串

语言能力

13. (L) 说："告诉我所有您喜欢吃的东西"，和/或"告诉我所有您喜欢在早饭/晚饭/午饭时做/吃的东西"，在 1 分钟内记录。

□ 2　说出 4 样或更多的东西

□ 1　说出 1 样、2 样或 3 样东西

记忆力

14. (M) 说："您还记得我的名字吗?"

说："(是的)，我的名字是________。"

测试这道题时采用与前面所说的完全相同的名字或称谓。

□ 2　自动说出正确的答案

□ 1　说出接近正确的答案，如将 "凯伦"说成"卡罗"，或将"史密斯先生/太太"说成"史密特先生/太太"等

语言能力

15. (L) 向受试者展示茶杯的照片，并说："这是什么?"

□ 2　"茶杯"

□ 1　说出与之接近的词汇，如"杯子"或"玻璃杯"

应用能力

16. (PR) 说："告诉我您是怎样使用这样东西的。"

□ 2　向测试者清楚地示范该物品的使用方法

□ 1　做出接近正确的表示，如受试者将手抬了起来，却没有明确地使之凑近受试者自己的嘴

语言能力

17. (L) 如果受试者在第 15 个问题中得了 2 分，则此题可给 2 分，但前提是必须完成此题，以便于以后测试其回忆能力。

说："拿住这样东西。"(把杯子给受试者)"(再问)这是什么?"

□ 2　自动说出正确的答案，或者患者已经正确地回答了第 15 个问题

□ 1　说出接近正确的答案

应用能力

18. (PR) 让受试者拿住杯子，同时说："再向我演示你是如何使用这样东西的?"

□ 2　向测试者清楚地示范该物品的使用方法

□ 1　做出接近正确的表示，如受试者将茶杯举了起来，却没有明确地使之凑近受试者自己的嘴

语言能力

19. (L) 如果受试者在第 15 题或第 17 题回答正确，则可跳过此题，并给予满分(1 分)。

说："这是一顶帽子还是一个茶杯?"

□ 1　"杯子"，或受试者已经正确地回答了第 15 题或第 17 题

□ 0　"帽子"

(此题没有可得 2 分的答案)

说："我希望您记住这只茶杯。"(拿起茶杯)，说："请尽量记忆，因为我将要在几分钟后向你提出与此有关的问题。"

20. (L) 向受试者展示勺子的照片，说："这是什么?"

□ 2　"勺子"

□ 1　说出与之接近的答案，如"银器/铜器"

应用能力

21. (PR) 说："告诉我您是怎样使用这样东西的。"

□ 2　向测试者清楚地示范该物品的使用方法

□ 1　做出接近正确的表示，如受试者将勺子举到自己的嘴边，却不把嘴凑上去

(续表)

语言能力

22. (L) 如果受试者在第 20 题中已经得到了 2 分，则此题可给 2 分，但前提是必须完成此题，以便于以后测试其回忆能力。

说："拿住这样东西。"(把勺子给受试者)再问："这是什么？"

□ 2 自动说出正确的答案，或者患者已经正确地回答了第 20 个问题

□ 1 说出接近正确的答案，如"银器/铜器"

应用能力

23. (PR) 让受试者拿住勺子，同时说："再向我演示你是如何使用这样东西的？"

□ 2 向测试者清楚地演示该物品的使用方法

□ 1 做出接近正确的表示，如受试者将勺子举起来，却没有将其凑近自己的嘴

语言能力

24. (L) 如果受试者在第 22 题或第 20 题中回答正确，则可跳过此题，并给予满分(1 分)。

说："这是一只靴子还是一个勺子？"

□ 1 "勺子"，或受试者已经正确地回答了第 20 题或第 22 题

□ 0 "靴子"

(此题没有可得 2 分的答案)

再次向受试者展示茶杯和勺子，并说："我希望您记住这把勺子"(拿起勺子)，"还有这个茶杯"(拿起茶杯)，"因为我将要在几分钟后向你提出与此有关的问题，仔细看一下并尽量记住。"

记忆力

25. (M) 把茶杯放在白板上，同时按照下面的顺序再放上两样其他的东西：

检查者的左侧	中央	检查者的右侧
塑料容器	盘子	**茶杯**

说："这里面哪个(项目/物品/东西)是我刚才请您记住的？"

把勺子放在白板上，同时按照下面的顺序再放上两样其他的东西：

检查者的左侧	中央	检查者的右侧
勺子	铲子	叉子

说："这里面哪个(项目/物品/东西)也是我刚才请您记住的？"

□ 2 说出"茶杯"和"勺子"

□ 1 要么说出了"茶杯"，要么说出了"勺子"

再次向受试者展示茶杯和勺子，并说："我希望您记住这把勺子"(拿起勺子)，"还有这个茶杯"(拿起茶杯)，"因为我将要在几分钟后向你提出与此有关的问题，仔细看一下并尽量记住。"

语言能力

26. (L) 向受试者展示一个蓝色的木块说："这是什么颜色的？"

如果受试者没有反应，则可提示说："这是蓝色的还是红色的？"

□ 2 自动说出正确的答案

□ 1 受试者说出一种接近正确的颜色(如紫色、海蓝色等)，或者受试者从给定的选择答案中选出了正确的颜色

视空间能力

27. (VS) 把蓝色、绿色和红色的木块按照下面的顺序分别放在白板上：

检查者的左侧	中央	检查者的右侧
蓝色	绿色	红色

拿着一个蓝色木块在受试者面前来回移动，以引导受试者看这个木块，说："哪个木块(手指着白板或轻拍桌子)和我手里的颜色相同？"

如果受试者没有反应，则可提示说："这是我的蓝色木块，出示你的蓝色木块。"(手指着测试者手里的蓝色木块和白板上的木块)

如果受试者的回答不正确或者没有反应，则拿起蓝色的木块，说："是这个，就是这个木块。"

（续表）

☐ 2　自动说出正确的答案
☐ 1　在提示后说出正确的答案
☐ 0　由检查者说出正确的木块

记忆力

28.（M）改变木块的摆放顺序如下：

检查者的左侧	中央	检查者的右侧
绿色	蓝色	红色

说："把那块木块还给我——也就是你刚才给过我的同一个木块(我给你看过的)。"

如果受试者没有反应，则可提示说："哪一个是你刚才给过我的木块(也就是我给你看过的木块)？是这一块吗？是这块还是那一块？"(手指着白板)。

如果受试者的回答不正确或者没有反应，则拿起蓝色的木块，说："是这个，就是这个木块。"

☐ 2　自动说出正确的答案
☐ 1　在提示后说出正确的答案
☐ 0　由检查者说出正确的木块

视空间能力

29.（VS）说："现在给我一块不同的木块，要不同于刚才我给你看的那块木块。"

如果受试者没有反应，则提示说："这是一块蓝色的木块。"(拿起蓝色的木块)，说："再给我一块不同颜色的木块。"

☐ 2　自动做出正确的反应
☐ 1　在提示后做出正确的反应

语言能力

30.（L）

a）向受试者展示红色的木块说："这块木块是什么颜色的？"

如果受试者没有反应，则提示说："这是蓝色的还是红色的？"

☐ 2　自动说出正确的答案
☐ 1　受试者说出与正确答案相近的颜色(如粉色或橘黄色)，或者受试者从给定的选择答案中选出了正确的答案

b）向受试者展示绿色的木块说："这块木块是什么颜色的？"

如果受试者没有反应，则提示说："这是蓝色的还是绿色的？"

☐ 2　自动说出正确的答案
☐ 1　受试者说出与正确答案相近的颜色(如橄榄色或柠檬色)，或者受试者从给定的选择答案中选出了正确的答案

c）向受试者展示黑色的方形木块说："这是什么形状的？"

如果受试者没有反应，则提示说："这是方形的还是圆形的？"

☐ 2　自动说出正确的答案
☐ 1　在提示后才说出正确的答案

视空间能力

31.（VS）把黑色的各种形状的木块按照下面的顺序分别放在白板上：

检查者的左侧	中央	检查者的右侧
三角形	圆形	**方形**

拿起一个形状类似的黑色方形木块，把该木块在受试者面前来回移动，以引导受试者注视这个木块，说："这些木块中哪一块的形状与这个木块相同？"(说时以手势示意白板或者用手轻轻拍打桌面)。

如果受试者没有反应，则可提示说："这是一块方形的木块，请你也向我展示一块方形的木块。"(可辅以清楚的手势示意)。

如果受试者仍然没有反应或者没有拿起正确的木块，则说："是这块，这就是方形的木块。"

☐ 2　自动说出正确的答案
☐ 1　在提示后才说出正确的答案
☐ 0　回答不正确或者由检查者自己拿起了正确的木块

（续表）

记忆力

32. (M) 按照下面的顺序重新摆放白板上的木块：

检查者的左侧	中央	检查者的右侧
圆形	方形	三角形

说："把那块木块还给我——和您刚才给我的木块相同(也就是我给你看过的那块木块)。"

如果受试者没有反应，则可提示说："哪一块是你刚才给过我的木块(也就是我给你看过的木块)？是这一块吗？是这块还是那一块？"(手指着白板上的木块)。

如果受试者仍然没有反应或者没有拿起正确的木块，则说："是这块，就是这块木块。"

☐ 2 自动说出正确的答案
☐ 1 在提示后才说出正确的答案
☐ 0 回答不正确或者由检查者自己拿起了正确的木块

视空间能力

33. (VS) 说："现在递给我一块不同形状的木块，要和我刚才给你的木块形状不同。"

如果受试者没有反应，则提示说："这是一块方形的木块"(拿起方形的木块)，说："再给我一块形状不同的木块。"

☐ 2 自动说出正确的答案
☐ 1 在提示后才说出正确的答案

语言能力

34. (L)

a) 向受试者展示一块圆形的木块说："这是什么形状的？"

如果受试者没有反应，则提示说："这是方形的还是圆形的？"

☐ 2 自动说出正确的答案(答"圆形"或"环形"都可以)
☐ 1 在提示后才说出正确的答案

b) 向受试者展示一块三角形的木块说："这是什么形状的？"

如果受试者没有反应，则提示说："这是方形的还是三角形的？"

☐ 2 自动说出正确的答案
☐ 1 在提示后才说出正确的答案，或者答道"锥形"

结构能力

35. (C)

a) 说："画一个圆圈。"

如果受试者没有反应，可为其先画一个圆圈作为示范，并说："照着这个画。"

☐ 2 自动画出环形、椭圆形或卵圆形的图案(允许因小的疏忽而画得形状不规范)
☐ 1 受试者画出接近正确的图案。如一个至少含有半圆的形状，或者在测试者的提示下画出正确的图案，或者在测试者画出的圆圈上描画
☐ 0 直线，点，等等。

b) 说："画一个正方形。"

如果受试者没有反应，可为其先画一个正方形作为示范，并说："照着这个画。"

☐ 2 受试者画出正方形、四边形或者长方形(允许因小的疏忽而画得形状不规范)
☐ 1 受试者画出接近正确的图案。如图形的一角没有闭合，但是若闭合就可构成一个正方形(但不能是三角形)，或者在测试者的提示下画出正确的图案，或者在测试者画出的正方形上描画
☐ 0 直线，点，等等。

注意力

36. (ATT) 说："我要拍打这个桌子，请计数我拍打桌子的次数。现在开始，仔细听！"

拍三下桌子，每次拍打的时间应比 1 秒钟稍短些，同时口中数着"1-2-3"，说："现在请你数数我拍桌子的次数，请你一直跟着数下去，不要中断"，拍 5 下桌子。本题只能提示一次。

☐ 2 受试者无须提示即可自己数出测试者 5 次拍击桌面
☐ 1 在测试者的提醒下，受试者数出 5 次
☐ 0 受试者在测试者提醒 1 次以上的情况下才数出 5 次，或者根本没有数出 5 次

（续表）

37.（ATT）勾起你的手指，以引起受试者的注意，说："看着我的手指，我竖起了3个手指"，测试者竖起第一、第二和第三个手指。然后，扳起大拇指，说："现在，我竖起了一个手指。"然后，扳起大拇指和无名指说："现在，请您数数我的手指""（对），是两个手指。"

然后，只竖起大拇指，如果受试者没有自发地数测试者的手指，测试者就要说："我希望您来数数我的手指，就这样一直数下去，不要停。"

在整个测试过程中，测试者只能提醒受试者一次。按照下面的顺序扳起相应的手指：

大拇指和无名指	大拇指	大拇指、示指和中指	无名指	所有上述4个手指
☐	☐	☐	☐	☐

☐ 2　如果受试者在测试者5次展示自己的手指时都能正确的数出来，且不中途停顿

☐ 1　如果受试者在测试者5次展示自己的手指时都能正确的数出来，但中途曾停顿过1次，且受到了测试者的1次提醒

☐ 0　如果数得不对或者受试者需要接受1次以上的提醒才能继续下去，完成计数

记忆力

38.（M）把茶杯放在白板上，同时按照下面的顺序再放上两样其他的东西：

检查者的左侧	中央	检查者的右侧
量杯	**茶杯**	碗

说："这里面哪个（项目/物品/东西）是我刚才请您记住的？"

拿掉所有的三样东西，把勺子放在白板上，同时按下面的顺序放上两样其他的东西：

检查者的左侧	中央	检查者的右侧
小刀	量勺	**勺子**

说："这里面哪样东西也是我刚才请您记住的？请指出来。"

☐ 2　说出"茶杯"和"勺子"

☐ 1　要么说出了"茶杯"，要么说出了"勺子"

到此为止，正式的"面对面"测试已经结束了，而测试者应该告诉受试者，他们可以准备离开了。

对名字的定向力

39.（ON）在受试者走回候诊室的过程中或在其准备离开的过程中，测试者站在受试者的正后方，并呼唤他/她的名字。

☐ 2　自发地做出正常反应，即受试者转过身来

☐ 1　有一定的反应（受试者做出语音的或非语音的反应，但其似乎对声音传来的方向不甚确定）

☐ 0　没有反应

语言能力

40.（L）如果受试者对第39题有反应，则测试者可吸引受试者与自己对话，说："你觉得怎么样？"

如果受试者只回答一个字或词（如"好"、"不错"），则鼓励其再做更多的反应，说："你这个周末有什么计划？""今天有人会来拜访您吗？"

当受试者对第39题没有反应时，测试者即可在受试者离开前的任何时间向其询问上述（那些）问题。

☐ 2　受试者连贯而恰当地回答了测试者所提出的1个或多个问题，所回答的内容必须为完整的句子

☐ 1　受试者对测试者的问题给予恰当的回答，但所答内容并非完整的句子。如"好"，或只有2到3个词，如"我还不错。"或"对，我还行。"

SIB已经有各种不同语言版本，如德文、意大利文、希腊文、挪威文、韩文、中文，中文版亦有比较好的信度与效度。

Saxton等于2005年发表SIB简短版，有9个项目：语言表达、言语记忆、

非言语记忆、社会交往、颜色命名、运用、阅读和书写、流畅性及注意力，耗时10～15 min。Schmitt 等于 2013 年提出一个由 8 个项目组成的 SIB-8 版本，项目包括何年、何月、写名字、句子回忆、流畅性、命名调羹、使用调羹、数字广度，耗时 3 min，每个项目 0～2 分，满分 16 分。SIB-8 与 SIB 全版的相关系数是0.859，与 MMSE 总分的相关系数是 0.716。

Malloy(2010)评估 22 名、平均 MMSE＝14 分的高龄老人的认知功能与活动能力，发现 SIB 与工具性功能活动评估量表(assessment of instrumental function, AIF)的相关性为 0.664，其 SIB 因子分中，相关性从高到低依次是：语言、定向力、记忆、运用能力、空间结构、视知觉、注意力、呼名回应，后 4 个因子的相关性无显著性。社会交往得分因天花板效应而无法进行相关分析。

在美金刚的多项临床研究中均应用严重损害量表(SIB)作为主要的疗效指标。在两项美金刚治疗中重度 AD 患者的荟萃分析中，研究结果显示美金刚较安慰剂可显著改善 SIB 评分(Patrizia Mecocci，2009 和 Murat Emre，2008)。在另外两项随机、双盲、安慰剂的对照研究中，美金刚均较安慰剂显著改善 SIB 评分(Reisberg 等，2003 和 Tariot 等，2004)。

有的研究认为 SIB 仅适用于重度 AD，由于天花板效应，SIB 对于有效区分轻度与中度 AD 不够敏感，故 Peavy 等于 1996 年编制了严重认知损害量表(SCIP)，包括注重、语言、记忆、运动速度和协调性、知觉推理、视觉空间、计算及行为举止等 8 个子因子，耗时约 30 min。总分 245 分，其中语言分为 88 分。在 41 例 MMSE 平均分为 8 分的 AD 患者中，SCIP 平均 170 分，SIB 平均 65 分，DRS 平均 58 分。

（周 燕 郭起浩）

第十二章 自评量表

第一节　常用情绪自评量表

一、流调用抑郁自评量表(CES-D)

流调用抑郁自评量表(center for epidemiological survey depression scale,CES-D)有20个项目,4级评分,总分60分,由美国国立精神卫生研究院制定,可以问卷的形式进行问讯评估或自我测查(表12-1-1)。CES-D已被使用了20多年,通过定式精神问诊和明确的诊断标准证实了CES-D对NINDS卒中数据库患者的有效性,评分大于16分则高度提示临床抑郁(敏感性86%,特异性90%,预测阳性率达80%)。已证明其在老年卒中患者(观察者评估和自我评估)中与其他抑郁测量结果高度一致,也在多个研究中被用以评估卒中后抑郁症状。它还被用于心血管健康研究。

表12-1-1　流调用抑郁自评量表(CES-D)

项　　目	0分	1分	2分	3分
1. 我因一些小事而烦恼				
2. 我不大想吃东西,我的胃口不好				
3. 即使家属和朋友帮助我,我仍然无法摆脱心中的苦闷				
4. 我觉得我比不上一般的人				
5. 我在做事时无法集中自己的注意力				
6. 我感到情绪低沉				
7. 我感到做任何事都很费力				
8. 我觉得我的前途没有希望				

（续表）

项　　目	0分	1分	2分	3分
9. 我觉得我的生活是失败的				
10. 我感到害怕				
11. 我的睡眠情况不好				
12. 我感到高兴不起来				
13. 我比平时说话要少				
14. 我感到孤单				
15. 我觉得人们对我不太友好				
16. 我觉得生活得没有意思				
17. 我曾哭泣				
18. 我感到忧愁				
19. 我觉得人们不喜欢我				
20. 我觉得无法继续我的日常工作				

填写说明：以下是一些你可能有过或感觉到的情况或想法。请按照过去1周内你的实际情况或感觉，在适当的格子内划一"√"。
0分＝没有或几乎没有(过去1周内，出现这类情况的日子不超过1天)；
1分＝少有(过去1周内，有1至2天有过这类情况)；
2分＝常有(过去1周内，有3至4天有过这类情况)；
3分＝几乎一直有(过去1周内，有6至7天有过这类情况)。

在国内，CES－D总分的划界分是，≤15分为无抑郁症状；16～19分为可能有抑郁症状；≥20分为肯定有抑郁症状。

二、老年抑郁量表(GDS)

老年抑郁量表(the geriatric depression scale，GDS)是Yesavage和Brink于1983年编制的，它的主要优点是评分方法非常简洁，适合老人使用。GDS也是受试者自评的方法，若由评定员评估，通常总得分比自评低。有轻度认知障碍的老人仍然可以准确地完成GDS。GDS有30道判断题，满分30分，≤10分为无抑郁症状；11～20分为可能有抑郁症状；≥21分为肯定有抑郁症状。GDS有一个简短的版本是GDS－15，由15项组成。有的研究认为它比30项的GDS有更高的可接受性，得分判断标准是：0～4分正常；5～8分轻度抑郁；8～11分中度抑郁；12～15分重度抑郁。其他还有12项、10项、5项、4项和1项组成的版本。文献中常用的是15项与5项版本。表12－1－2中带＊的项目组成5项版本，带＃的组成15项版本。

表 12-1-2　老年抑郁量表(GDS)

序号	项　　目	回　答
1＊#	你对生活基本满意吗？	是　否
2#	你是否已放弃了许多爱好与兴趣？	是　否
3#	你是否觉得生活空虚？	是　否
4＊#	你是否感到厌倦？	是　否
5	你觉得未来有希望吗？	是　否
6	你是否因为脑子里一些想法摆脱不掉而烦恼？	是　否
7#	你是否大部分时间精力充沛？	是　否
8#	你是否害怕会有不幸的事落到你头上？	是　否
9#	你是否大部分时间感到幸福？	是　否
10＊#	你是否经常感到孤立无援？	是　否
11	你是否经常坐立不安，心烦意乱？	是　否
12＊#	你是否愿意待在家里而不愿去室外做些新鲜事？	是　否
13	你是否常常为将来的生活担心？	是　否
14#	你是否觉得记忆力比以前差？	是　否
15#	你觉得现在活着很开心吗？	是　否
16	你是否常感到心情沉重、郁闷？	是　否
17＊#	你是否觉得像现在这样活着毫无意义？	是　否
18	你是否总为已经过去的事忧愁？	是　否
19	你觉得生活很令人兴奋吗？	是　否
20	你开始一件新的规划很困难吗？	是　否
21#	你觉得生活充满活力吗？	是　否
22#	你是否觉得你的处境已毫无希望？	是　否
23#	你是否觉得大多数人比你强得多？	是　否
24	你是否常为一些小事伤心？	是　否
25	你是否常觉得想哭？	是　否
26	你集中精力有困难吗？	是　否
27	你早晨起来很快活吗？	是　否
28	你希望避开各种聚会吗？	是　否
29	你做决定很容易吗？	是　否
30	你的头脑像往常一样清晰吗？	是　否

填写说明：选择最切合您 1 周来感受的答案，在每题后选择“是”或“否”。

三、焦虑自评量表（SAS）

焦虑自评量表(self-rating anxiety scale SAS)由 Zung 编制(1971)。SAS

(表 12－1－3)采用 4 级评分，主要评定症状出现的频度，其标准为："1"表示没有或很少时间有；"2"表示有时有；"3"表示大部分时间有；"4"表示绝大部分或全部时间都有。20 个条目中有 15 项是用负性词陈述的，按上述 1～4 顺序评分。其余 5 项(第 5,9,13,17,19)注 * 号者，是用正性词陈述的，按 4～1顺序反向计分。分析指标：SAS的主要统计指标为总分。将 20 个项目的各个得分相加，即得粗分；用粗分乘以 1.25 以后取整数部分，就得到标准分，或者可以查表做相同的转换。Zung 根据美国受试者测评结果，规定 SAS 的标准分 50 分作为焦虑症状分界值。全国协作组吴文源等人对 1 158 例正常人(常模)测评结果分析，正评题 15 项平均值 1.29±0.98；反向 5 项均分 2.08±1.71；20 项总分均值 29.78±0.46，可作为代表常模总分均值之上限。

表 12－1－3　焦虑自评量表 (SAS)

项　　目	1	2	3	4	他评
1. 我觉得比平常容易紧张或着急	1	2	3	4	1□
2. 我无缘无故地感到害怕	1	2	3	4	2□
3. 我容易心里烦乱或觉得惊恐	1	2	3	4	3□
4. 我觉得我可能将要发疯	1	2	3	4	4□
*5. 我觉得一切都很好，也不会发生什么不幸	4	3	2	1	5□
6. 我手脚发抖打颤	1	2	3	4	6□
7. 我因为头痛、颈痛和背痛而苦恼	1	2	3	4	7□
8. 我感觉容易衰弱和疲乏	1	2	3	4	8□
*9. 我觉得心平气和，并且容易安静坐着	4	3	2	1	9□
10. 我觉得心跳得很快	1	2	3	4	10□
11. 我因为一阵阵头晕而苦恼	1	2	3	4	11□
12. 我有晕倒发作，或觉得要晕倒似的	1	2	3	4	12□
*13. 我吸气、呼气都感到很容易	4	3	2	1	13□
14. 我的手脚麻木和刺痛	1	2	3	4	14□
15. 我因为胃痛和消化不良而苦恼	1	2	3	4	15□
16. 我常常要小便	1	2	3	4	16□
*17. 我的手脚常常是干燥温暖的	4	3	2	1	17□
18. 我脸红发热	1	2	3	4	18□
*19. 我容易入睡并且一夜睡得很好	4	3	2	1	19□
20. 我做噩梦	1	2	3	4	20□

注：* 用正性词陈述。

四、淡漠评估量表(AES)

淡漠评估量表(apathy evaluation scale, AES)由 Marin 等编制(1991),通过对过去 4 周的表现进行评估,判定患者的淡漠程度。该量表包括临床医师版本(AES-C)、知情者版本(AES-I)(表 12-1-4)和患者版本(AES-S)(表 12-1-5)。每个版本均包含相同的 18 道题,每一道题均采用 4 级评分,主要评定症状同意的程度,其标准为:"0"表示完全不是;"1"表示有点同意;"2"表示基本同意;"3"表示完全同意。分析指标:AES 的主要统计指标为总分,得分越高,提示淡漠症状越严重,目前并未明确规定淡漠症状的分界值。研究显示,这 3 个版本的 Cronbach α 系数为 0.76～0.94,3 个版本之间的相关性 r=0.43～0.72,患者版本淡漠与抑郁的相关性 r=0.43,知情者版本淡漠与抑郁的相关性 r=0.27,临床医师版本淡漠与焦虑的相关性 r=0.35,患者版本淡漠与焦虑的相关性 r=0.42。

表 12-1-4 淡漠评估量表(AES-C 或 AES-I)

根据受试者过去 4 周的表现进行评估。

项　　目	完全不是	有点同意	基本同意	完全同意
1. 他/她对于各种事情有兴趣				
2. 他/她在计划的时间内做完事情				
3. 他/她认为自己主动去完成事情是重要的				
4. 他/她有兴趣接受新的体验				
5. 他/她有兴趣学习新的事情				
6. 他/她对任何事情都很努力				
7. 他/她对生活充满热情				
8. 他/她有始有终地做一件事情				
9. 对于有兴趣的事情,他/她愿意花时间去做				
10. 不需要每天告诉他/她应该做哪些事情				
11. 他/她关心周围各种事情				
12. 他/她有许多朋友				
13. 与朋友聚会,他/她觉得很重要				
14. 发生好事情时,他/她显得很高兴				
15. 他/她能够准确理解他/她的问题				
16. 对他/她来说,当天的事情当天做完是重要的				
17. 他/她有积极性和主动性				
18. 他/她会不断激励自己				

表 12-1-5 淡漠评估量表(AES-S)

根据你自己过去 4 周的表现进行评估。

项目	完全不是	有点同意	基本同意	完全同意
1. 我对于各种事情有兴趣				
2. 我能够在计划的时间内做完事情				
3. 我认为自己主动去完成事情是重要的				
4. 我有兴趣接受新的体验				
5. 我有兴趣学习新的事情				
6. 我对任何事情都很努力				
7. 我对生活充满热情				
8. 我有始有终地做一件事情				
9. 对于有兴趣的事情,我愿意花时间去做				
10. 不需要他人每天告诉我应该做哪些事情				
11. 我关心周围各种事情				
12. 我有许多朋友				
13. 与朋友聚会,我觉得这是很重要的				
14. 发生好事情时,我会很高兴				
15. 我能够准确理解我自己的问题				
16. 对我来说,当天的事情当天做完是重要的				
17. 我有积极性和主动性				
18. 我会不断激励自己				

五、多伦多移情问卷(TEQ)

多伦多移情问卷(the Toronto empathy questionnaire, TEQ)由 Spreng 等编制(2009)(表 12-1-6)。该量表包括知情者版本和患者版本。每个版本均包含相同的 16 道题,每一道题均采用 5 级评分,主要评定症状出现的频度,其标准为:"0"表示从不;"1"表示极少;"2"表示有时;"3"经常;"4"表示总是。16 道题目中有 8 项(第 1、3、5、6、8、9、13、16 题)是用正性词陈述的,按上述 0~4 顺序评分。其余 8 项(第 2、4、7、10、11、12、14、15 题)是用负性词陈述的,按4~1 顺序反向计分。分析指标:TEQ 的主要统计指标为总分。TEQ 得分与患者本身相关,如疾病严重程度($r=0.53$, $P=0.03$)、行为($r=-0.65$, $P=0.008$)、认知和情感等,知情者版本的得分还与其照料负担、与患者的关系情况等相关。

表 12-1-6　多伦多移情问卷(TEQ)

● **患者版本**

请认真阅读以下句子，根据你的感受或行为方式做出程度评估，圈出你的答案，尽可能反映你的真实情况。

项　　目	严	重	度		
1. 其他人觉得很兴奋时，我也会很兴奋	从不	极少	有时	经常	总是
2. 其他人的不幸遭遇不会扰乱我的心情	总是	经常	有时	极少	从不
3. 看到有人受到不公正的对待，我感到很失望	从不	极少	有时	经常	总是
4. 当表现很幸福的人靠近我，我不受影响	总是	经常	有时	极少	从不
5. 我让其他人高兴时，我自己也很喜悦	从不	极少	有时	经常	总是
6. 看到其他人没有我幸运时，我有亲切的、关心的感受	从不	极少	有时	经常	总是
7. 当有朋友谈他/她的烦恼时，我试图中断谈话，转移话题	总是	经常	有时	极少	从不
8. 其他人不开心时，即使口头没有表达出来，我也能够感受到并给予安慰	从不	极少	有时	经常	总是
9. 我发现我跟着其他人的情绪的步调走	从不	极少	有时	经常	总是
10. 有人患重病时，我并不同情	总是	经常	有时	极少	从不
11. 有人哭叫时，我会心烦意乱	总是	经常	有时	极少	从不
12. 我对其他人的感受一点也不感兴趣	总是	经常	有时	极少	从不
13. 当我看到有人苦恼的时候我极力鼓励帮助他/她	从不	极少	有时	经常	总是
14. 当我看到有人受到不公正的对待时，我并不同情他/她	总是	经常	有时	极少	从不
15. 当我听到有人说他/她很幸福时，我觉得这个人很无聊	总是	经常	有时	极少	从不
16. 当我看到有人受骗上当时，我愿意提醒保护他/她	从不	极少	有时	经常	总是

● **知情者版本**

请认真阅读以下句子，根据他/她的感受或行为方式做出程度评估，圈出你的答案，尽可能反映他/她的真实情况。

项　　目	严	重	度		
1. 其他人觉得很兴奋时，他/她也会很兴奋	从不	极少	有时	经常	总是
2. 其他人的不幸遭遇不会扰乱他/她的心情	总是	经常	有时	极少	从不
3. 看到有人受到不公正的对待他/她感到很失望	从不	极少	有时	经常	总是
4. 当表现很幸福的人靠近他/她，他/她不受影响	总是	经常	有时	极少	从不
5. 他/她让其他人高兴时，他/她自己也很喜悦	从不	极少	有时	经常	总是
6. 看到其他人没有他/她幸运时，他/她有亲切的、关心的感受	从不	极少	有时	经常	总是
7. 当有朋友谈他/她们自己的烦恼时，他/她试图中断谈话，转移话题	总是	经常	有时	极少	从不
8. 其他人不开心时，即使口头没有表达出来，他/她也能够感受到并给予安慰	从不	极少	有时	经常	总是
9. 他/她发现他/她跟着其他人的情绪的步调走	从不	极少	有时	经常	总是
10. 有人患重病时，他/她并不同情	总是	经常	有时	极少	从不

（续表）

项　　　目	严	重	度		
11. 有人哭叫时，他/她会心烦意乱	总是	经常	有时	极少	从不
12. 他/她对其他人的感受一点也不感兴趣	总是	经常	有时	极少	从不
13. 当他/她看到有人苦恼时，他/她极力鼓励帮助	从不	极少	有时	经常	总是
14. 当他/她看到有人受到不公正的对待，他/她并不同情	总是	经常	有时	极少	从不
15. 当他/她听到有人说他/她很幸福时，他/她觉得这个人很无聊	总是	经常	有时	极少	从不
16. 当他/她看到有人受骗上当时，他/她愿意提醒保护他/她	从不	极少	有时	经常	总是

（王　萍　郭起浩）

第二节　老年认知功能减退知情者问卷(IQCODE)

使用知情者问卷了解患者日常认知功能，是临床评估认知损害的方法之一。与客观神经心理检查方法比较，知情者问卷的优点是：① 与日常生活状态密切相关。受试者的日常生活表现不仅和认知有关，也和环境需要相关。一个活动范围有限的生活状态与一个需要较高智力水平参与的生活状态所需要的认知水平肯定是不同的。知情者问卷可以很好地了解到患者的功能状态与生活环境所需能力之间是否相匹配。② 易于接受。一些患者会因认知检查的目的是明确其认知有问题而感到不快。而知情者问卷则不需患者直接参与评估。③ 可用于不适于客观检查的患者。认知损害严重患者可能无法完成客观认知检查(即地板效应)，知情者问卷可以不受患者认知状态的影响。④ 知情者问卷不需面对面会晤，可以通过信件或电话完成。⑤ 适于纵向观察。认知检查方法通常只能得到一个当前的得分，无法得知与患者既往情况的差异。知情者问卷则可明确患者的日常功能较前有无改变。⑥ 无跨文化差异。认知检查方法通常无法克服文化差异问题。知情者问卷涉及的多是有共性的日常生活内容(如想起熟人的名字)，适于不同文化背景患者应用。另外，由于知情者仅需比较患者目前与既往的生活情况，其评估结果也很少受到受教育程度的影响，可在不同的社会经济境况下使用。

当然，知情者问卷也有其自身的不足：首先就是患者有时会缺少一个合适的知情者，无法使用知情者问卷。其次，认知测验可通过一些方法评估患者

某一认知域的能力，如记忆、执行功能。但知情者问卷的内容为日常认知能力，混杂了多种认知功能，因而它不适用于评估某一特定认知功能。此外，知情者的情绪状态和人格特征、与受试者的关系融洽与否也都会干扰评估结果。

一、IQCODE 的内容

老年认知功能减退知情者问卷(informant questionnaire on cognitive decline in the elderly, IQCODE)是目前最常使用的知情者问卷之一，由 Jorm 和 Jacomb 于 1989 年在对澳大利亚痴呆人群的研究实践中发展而来。目的是要建立一个不受教育程度、文化背景、病前能力因素影响的评估问卷。该问卷不同于以往的量表，更加注重测量 10 年期间老人的认知变化，而不仅仅是只对当前认知能力进行评估。之所以使用 10 年的时间历程，是当时的流行病学资料表明痴呆患者从有症状初起至死亡的病程不超过 10 年。最初的 IQCODE 版本包含 39 项问题，用于评价老年人日常认知功能的记忆(26 项)和智能(13 项)。记忆条目分为两大类：一是获得新知识的能力，另一是回忆已有知识的能力。在这两类条目中，记忆条目的设计涵盖了概念(knowledge of what)、过程(knowledge of how to)和视空间记忆功能。智能条目又进而分为两大类：一是词语方面(使用 WAIS 中的词语量表)，二是逻辑方面(使用 WAIS 的中的行为量表)。还有一小部分条目是全面评估记忆或智能改变的。后去除了 12 个知情者认为不易评估的条目和一个与总分相关性低的条目，剩余的 26 项具有更好的内在联系性和可施行性，组成了目前常用的 IQOCDE 量表。问卷所涉及的认知功能包括近期、远期记忆力，空间、时间定向力，计算力，学习能力及执行能力。IQCODE 评分方法将患者认知功能水平改变程度分为 5 个等级：1 为“好多了”；2 为“好一点”；3 为“没变化”；4 为“差一点”；5 为“差多了”。

测验时，请大声地读给受试者听：“我希望您能记起____先生(太太)10 年前的情形，来和他现在的情形相比较。10 年前是指 20____年。总共认识有____年。首先我要请教您____先生(太太)记忆力方面的情形，包括他对现在的日常生活和以前所发生的事情的记忆力。请记住，我们主要是比较____先生(太太)现在和他 10 年前的情况。所以，假如他在 10 年前就常常忘记东西放在哪里，而现在仍然如此，就请您回答没有什么变化。”

二、IQCODE 的不同版本

IQCODE 已被翻译为多种语言，包括汉语、芬兰语、德语、法语、荷兰语、意

大利语、日语、韩语、挪威语、波兰语、西班牙语、泰语等。其相应的版本可在 http://ageing.anu.edu.au/Iqcode 上免费下载使用。

(一) IQCODE 的版本

IQCODE 的版本可分为 4 种类型：长版(long form)、短版(short form)、极简版(very short form)和回顾版(retrospective form)。

1. 长版　即 26 项 IQCODE 版本，内容见表 12-2-1。

表 12-2-1　IQCODE 长版

项　　目	好多了	好一点	没变化	差一点	差多了	不知道(拒答)
1. 认得出家人和熟人的面孔	1	2	3	4	5	9
2. 记得家人和熟朋友的名字	1	2	3	4	5	9
3. 记得家人和熟人的职业、生日和住址	1	2	3	4	5	9
4. 记得最近发生的事情	1	2	3	4	5	9
5. 记得几天前谈话的内容	1	2	3	4	5	9
6. 话说到一半就忘记了要说什么	1	2	3	4	5	9
7. 记得自己的住址和电话号码	1	2	3	4	5	9
8. 记得今天是星期几、是几月份	1	2	3	4	5	9
9. 记得东西经常是放在什么地方	1	2	3	4	5	9
10. 东西未放回原位，仍能找得到	1	2	3	4	5	9
11. 能适应日常生活中的一些改变	1	2	3	4	5	9
12. 使用日常用具的能力(如电视机、铁锤等)	1	2	3	4	5	9
13. 学习使用新的家用工具与电器的能力	1	2	3	4	5	9
14. 学习新事物的能力	1	2	3	4	5	9
15. 能记住年轻及童年往事	1	2	3	4	5	9
16. 能记住年轻时所学的东西	1	2	3	4	5	9
17. 懂一些不常用的字	1	2	3	4	5	9
18. 看懂报纸杂志上的文章	1	2	3	4	5	9
19. 看懂电视或书本中讲的故事	1	2	3	4	5	9
20. 写信表达的能力	1	2	3	4	5	9
21. 知道一些重要的历史事件	1	2	3	4	5	9
22. 对日常生活事物自己会做决定	1	2	3	4	5	9
23. 会用钱买东西	1	2	3	4	5	9
24. 处理财务的能力(如退休金、到银行)	1	2	3	4	5	9
25. 处理日常生活上的计算问题(如知道要买多少食物，知道朋友或家人上一次来访有多久了)	1	2	3	4	5	9
26. 了解正在发生什么事件及其原因	1	2	3	4	5	9

2. 短版　指16项IQCODE版本，内容见表12-2-2。Jorm通过因子分析删除了长版中受教育程度影响较大、信度和效度欠佳的10项，形成了短版IQCODE。短版IQCODE被证实与长版有高度的相关性($r=0.98$)且在筛查痴呆方面有着同等的效能，而筛查时间则缩短了近1/3。此版本目前已有9国语言的翻译版，为大多数研究者所采用。

表12-2-2　IQCODE短版

项　　目	好多了	好一点	没变化	差一点	差多了	不知道(拒答)
1. 记得家人和熟人的职业、生日和住址	1	2	3	4	5	9
2. 记得最近发生的事情	1	2	3	4	5	9
3. 记得几天前谈话的内容	1	2	3	4	5	9
4. 记得自己的住址和电话号码	1	2	3	4	5	9
5. 记得今天是星期几、是几月份	1	2	3	4	5	9
6. 记得东西经常是放在什么地方	1	2	3	4	5	9
7. 东西未放回原位，仍能找得到	1	2	3	4	5	9
8. 使用日常用具的能力(如电视机、铁锤等)	1	2	3	4	5	9
9. 学习使用新的家用工具与电器的能力	1	2	3	4	5	9
10. 学习新事物的能力	1	2	3	4	5	9
11. 看懂电视或书本中讲的故事	1	2	3	4	5	9
12. 对日常生活事物自己会做决定	1	2	3	4	5	9
13. 会用钱买东西	1	2	3	4	5	9
14. 处理财务的能力(如退休金、到银行)	1	2	3	4	5	9
15. 处理日常生活上的计算问题(如知道要买多少食物，知道朋友或家人上一次来访有多久了)	1	2	3	4	5	9
16. 了解正在发生什么事件及其原因	1	2	3	4	5	9

3. 极简版　Ehrensperger将16项IQCODE德语版进一步缩减为7项，并证实7项版与16项版在区分MCI与轻度AD上几乎有相同的效能。

4. 回顾版　应用于突发事件导致的患者认知改变的评价，如死亡或突发急性疾病。回顾版IQCODE可应用于评价生前未收集到足够认知功能信息的捐献大脑者，在纵向研究中用于评价随访期间死亡患者的认知情况，以及评价患者在卒中、谵妄，或进入重症监护病房之前的认知状况。

(二) 优化IQCODE

以上版本均得到IQCODE作者Anthony F. Jorm的认可。此外，还有许多研究者尝试在应用中继续优化IQCODE，主要集中在以下两个方面。

1. 不同的追溯时间　一些使用者发现很难找到可以陪伴受试者达10年之久的知情者。此外患者超过10年以上的功能改变很可能是由多种因素如退休等引起的，而不单纯是由于认知功能的改变；而且随着对痴呆早期诊断和治疗的进展，短期的知情者评价可以更有利于监测痴呆症状的出现。Barba和Pisani改为5年时限，Patel根据知情者与个体的关系采用更灵活的时间限制。目前尚没有报道这些改动是否会影响到效度。另一些作者尝试进一步减少到2年的可能。Schultz请日间护理人员评估他们照顾的精神发育迟滞老年个体，结果发现信度和效度都仅是中等。而极简版IQCODE的作者Ehrensperger也是使用2年时限的IQCODE来区分MCI和轻度AD，正确诊断率分别可达79.9%和90.7%。

2. 其他版本项目　Butt研究发现IQCODE量表各项问题在筛查痴呆时的权重不同，因此存在根据不同诊断需要而简化量表的可能。Morales等建立了17项短版的西班牙语IQCODE(short Spanish version of IQCODE，SS-IQCODE)，其中12项与英文版的短版相同。SS-IQCODE在预测痴呆、不受受教育程度及病前状态影响方面与26项长版类似。Fuh等使用判别分析方法选择了26项中文长版中的17项组成了新的短版IQCODE，并发现其效度与长版一致。通过对399名社区居民和61名痴呆患者进行调查，发现仅单独用其中的2项问题(回忆几天前的谈话内容；处理财务的能力)就可以在筛查痴呆时有较高的准确度。Senanarong等使用泰语版IQCODE进行logistic回归分析，结果提示3项问题对于痴呆有显著的辨别作用（记得今天是几月份，星期几；学习使用新的家用工具和电器的能力；处理日常生活中的计算问题)。单独使用此3项筛查痴呆可达到敏感度85%，特异度92%。

三、评分方法

评分方法分为两种：一种是求得所有项目的平均分(总分除以项目数)，得分为1～5；另一种是直接将所有项目的得分相加，得到一个总分：26项版为26～130，16项版为16～80。一般认为，26项版允许3个缺失项(即填9的数目)，16项版仅允许2个。但也有其他研究允许5～6个缺失项的。

四、常模、信度和效度

1. 常模　不同版本及人群IQCODE辨别痴呆的划界分(cut off point)有

很大变异性。但总体来讲，社区人群的划界分为3.3～3.6，住院患者较社区人群偏高，为3.4～4.0，而短版划界分值会高于长版。

2. 内部一致性与重测信度　IQCODE量表具有良好的内部一致性，在涵盖了不同人群和语言环境的多项研究中IQCODE量表的Cronbach's α系数均很高，为0.93～0.97。重测信度Pearson相关系数在3天时达到0.96，在1年时达到0.75；国内研究结果为IQCODE量表的内部一致性系数达0.83，重测信度Pearson相关系数为0.86。

3. 效度　针对社区人群，IQCODE对痴呆的敏感度为76%～89%，特异度为65%～96%，其中中文版本的IQCODE量表的敏感度和特异度分别为89%和88%。在医院或诊所等以患者为研究对象时，IQCODE对痴呆的敏感度可达74%～100%，特异度达71%～94%。

Flicker注意到IQCODE在评价住院老年患者和社区正常老年人的阳性预测值具有差异性，分别为0.85和0.45，这种差异性说明了IQCODE的效度和其他筛查量表一样，均受到应用人群特性的影响。

五、临床评价

（一）IQCODE的影响因素

1. 患者方面　相对于客观的认知功能检查量表，IQCODE的优势在于既不受患者年龄、受教育程度影响，也不受发病前智能状态的影响。

运用多种衡量教育程度的方法（如受教育年限，离开学校年龄，成就水平等），多国的研究均证明IQCODE分值与患者受教育程度没有相关性（相关系数为−0.02～0.07），且在这些研究中多支持IQCODE分值不受患者的年龄和性别的影响。

有研究使用NART或其他阅读测试与IQCODE得分比较，发现两者没有显著相关性，说明IQCODE得分不受发病前智能水平影响。

同时还有一些探讨IQCODE得分与患者焦虑、抑郁和精神应激状态关系的研究，大多数发现它们之间有一定的弱相关性，这可能是由于抑郁状态可能会影响一定程度的认知功能，但也可能是知情者难以区分焦虑抑郁与认知功能下降所致。

2. 知情者方面　IQCODE分值不受知情者与患者的关系类型，或关系年限，或知情者的年龄及文化程度的影响。大多数研究表明，知情者的情绪状态，如焦虑、抑郁及因照料患者产生的压力会影响IQCODE分值。与MMSE

相比，知情者的焦虑及抑郁状态对 IQCODE 分值有更强的影响：知情者焦虑状态与量表得分相关性为 $r=0.23$(IQCODE) vs. $r=-0.05$(MMSE)；知情者抑郁状态与量表得分的相关性为 $r=0.22$(IQCODE) vs. $r=0.14$(MMSE)。

此外，研究人群的不同对 IQCODE 分值也会产生影响。医院人群调查时，由于照料者往往有焦虑情绪，以及试图获得更多的医疗关注，他们往往会夸大认知下降水平，而在社区调查中，家属往往有不愿接受亲人被诊断痴呆的情绪，因而 IQCODE 分值会较低。

（二）IQCODE 与 MMSE 联合使用

两种量表的联合应用可能提高筛查痴呆的性能。

最早发表的进行两种量表联合应用的来自瑞士一家老年医院和记忆诊所的 116 名患者。Mackinnon 和 Mulligan 使用三种 IQCODE 和 MMSE 联合应用方法：①"Or"法：若任一测试阳性则认为患者测试结果为阳性，其敏感度 93%，特异度 81%。②"And"法：只有当两项测试均为阳性使患者的测试结果为阳性，其敏感度 86%，特异度 85%。③"weighted sum"法：用 logistic 回归计算出两种测试结合后的最佳预测值，并建立相应的方程来检测阳性患者，敏感度 97%，特异度 85%。说明使用"And"法并不比单独使用 MMSE 优越；使用"Or"法可提高敏感度，而特异度没有明显降低；"weighted sum"法的效度最好，但在临床上应用比较不便利。在社区人群中也有类似的结论。

Narasimhalu 等发现"weighted sum"法有最优的 AUC 面积和特异性，可以使文化程度对 MMSE 引起的偏差减小，因此对于亚洲老年人群的研究尤其有意义。而"Or"法则有最佳的敏感性。

（李 放）

第三节 多因素记忆问卷(MMQ)

2002 年，Troyer 等根据成人元记忆问卷（metamemory in adulthood questionnaire，MAQ，Dixon，1998）和记忆功能问卷（memory functioning questionnaire，MFQ，Gilewski，1990）在内的 8 种问卷内容和条目，通过信度和效度经验，编制多因素记忆问卷（multifactorial memory questionnaire，MMQ），包括记忆满意度、记忆能力和记忆策略三部分，共 57 个项目，每个

项目分5个等级：非常同意为4、同意为3、不确定为2、不同意为1、完全不同意为0。该版本内容全面，有优良的内容效度、结构效度、再测信度和一致性信度，已经被翻译成不同语言版本使用。笔者将其翻译为中文，个别文化相关的项目做了修改并增加了3项，每个因子20项，共60项，见表12-3-1。

表12-3-1　多因素记忆问卷(MMQ)

(一)请你根据最近2周的情况做出回答

1. 我对自己的记忆力还算满意
非常同意—0　同意—1　不确定—2
不同意—3　完全不同意—4

2. 我的记忆力糟透了
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

3. 如果是重要的事情，我一般能记住
非常同意—0　同意—1　不确定—2
不同意—3　完全不同意—4

4. 当我出现遗忘的时候，我害怕有严重的记忆问题，比如可能患了老年痴呆
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

5. 我的记忆力比其他同龄人要差
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

6. 我相信自己能记住要记的事情
非常同意—0　同意—1　不确定—2
不同意—3　完全不同意—4

7. 一想到自己的记忆能力，我就不开心
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

8. 我担心其他人发现我的记忆力不是很好
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

9. 当我为记住某些事情烦恼时，我不勉强自己去记
非常同意—0　同意—1　不确定—2
不同意—3　完全不同意—4

10. 我很关注自己的记忆力
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

11. 最近我的记忆力确实在走下坡路
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

12. 一般来说，我对自己的记忆力还算满意
非常同意—0　同意—1　不确定—2
不同意—3　完全不同意—4

13. 当我有什么事记不住时，我没什么不开心的
非常同意—0　同意—1　不确定—2
不同意—3　完全不同意—4

14. 我担心会忘记一些重要的事情
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

15. 我对自己的记忆力很灰心
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

16. 当我出现健忘时，我很苦恼
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

17. 就我的年龄而言，我对自己的记忆力很满意
非常同意—0　同意—1　不确定—2
不同意—3　完全不同意—4

18. 我担心自己的记忆能力
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

19. 与其他年龄差不多的老人相比，我的记忆还不错
非常同意—0　同意—1　不确定—2
不同意—3　完全不同意—4

20. 想到我的记忆问题，我的睡眠也受到影响
非常同意—4　同意—3　不确定—2
不同意—1　完全不同意—0

（续表）

（二）以下情况在过去 2 周出现的频率

1. 忘记按时还钱
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

2. 将日常用品比如钥匙、眼镜之类的东西摆错地方
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

3. 忘记刚刚看过的电话号码
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

4. 忘记刚才碰到的人的姓名
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

5. 忘记本要带走的东西
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

6. 忘记约定的时间
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

7. 忘记原本打算要做的事情，举个例子，你走进房间，但是忘记进去后要做些什么
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

8. 忘记出去办事
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

9. 要用某个词表达时，就是想不起来
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

10. 记不清今天早上看的报纸或杂志文章的细节
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

11. 忘记服药
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

12. 不能回忆早就记住的人名
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

13. 忘记带口信
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

14. 在交谈中忘记将要说的话
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

15. 忘记过去记得很清楚的生日或纪念日
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

16. 忘记常用的电话号码
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

17. 由于健忘，重复向某人讲同样的故事或笑话
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

18. 记不清前几天把东西放哪了
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

19. 忘记买本来打算买的东西
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

20. 忘记近期与人交谈的细节
不出现—4　很少出现—3　有时出现—2　经常出现—1　一直出现—0

（三）以下改善记忆的方法，你在过去 2 周使用的频率

1. 用计时器或闹钟来提醒自己做某件事
一直使用—4　经常使用—3　有时使用—2　很少使用—1　从不使用—0

2. 求助其他人（比如家庭成员）提醒你记住一些事情或做事情
一直使用—4　经常使用—3　有时使用—2　很少使用—1　从不使用—0

（续表）

（三）以下改善记忆的方法，你在过去2周使用的频率

3. 将常用的地址、电话等信息写在卡片上随身带着
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

4. 借助视觉形象来加深记忆，比方说，记住某个人的相貌，从而来记住人名
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

5. 在日历上做一些备注，比如标明各种活动时间
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

6. 对人名或要记的各种名称做拆字、谐音、字谜、编制口诀之类的方式加强记忆
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

7. 对你想记的内容做出分类，比如根据肉类、蔬菜、饮料、调料等不同类别列出采购清单
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

8. 大声说出某样东西来加深记忆，比如大声说出你看到的电话号码
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

9. 养成一些习惯来记住一些重要的事情，比如你出门时常常检查一下是否带了钱包和钥匙
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

10. 列出清单或表格，比如购物清单或要做的事情的清单
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

11. 特别花心思记一些事情，比方说通过记住一些关键细节来加深记忆
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

12. 把东西放在醒目的地方来提醒自己做某事，比如把伞放在门前，这样出门时就会记得带上了
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

13. 每隔一段时间想一想，通过反复回忆来加深记忆
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

14. 把要记的信息编进故事里来加深记忆
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

15. 在本子上做笔记来加深记忆
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

16. 首字归纳来加深记忆，比如，用“文新”记“文汇报、新民晚报”
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

17. 对要记的内容特别集中注意力来加深记忆
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

18. 给自己写注意事项或备忘录（并不是指在日历上或者笔记本上简单标记一下）
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

19. 试着回忆一些过去的事情来整理记忆，比如说通过回忆记起将东西错放在哪里了
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

20. 为了提高记忆力而听音乐、背诗词、练书法、画画等
一直使用—4　经常使用—3　有时使用—2
很少使用—1　从不使用—0

笔者在正常老人、MCI和轻度AD中试用MMQ中文版，发现轻度AD中自评是不准确的，即自评得分与客观的神经心理测验没有相关性，所以，MMQ主要是正常老人进行认识训练时作为自我评价的手段，能否作为MCI的检测

方法尚需更多研究，估计自评与家属评估相互结合填写可以克服不准确的问题。除了对老人观察了解不足，出于“家丑不可外扬”的社会心理因素，家属未必愿意公开指出受试者的记忆衰退；也可能出于对长辈的尊重、奉承或利益冲突的原因，故意出现偏差，所以，家属填写结果与客体记忆的相关性也不一定很高。

（郭起浩）

第四节 痴呆知情者问卷（AD8、OLD、AQ）

AD8 是 8 项知情者半结构性晤谈量表，是一项询问知情者的认知损害筛查工具。华盛顿大学 Galvin 等于 2002 年根据文献归纳出 55 个问卷项目，经过 290 例验证，于 2005 年发表正式版本（表 12-4-1），共 8 个条目，用于评估患者因认知问题导致的改变，耗时小于 2 min。

表 12-4-1 痴呆知情者问卷（AD8）量表

项　　目	是	不是	无法判断
1. 判断力出现问题（解决日常生活问题、经济问题有困难，如不会算账了；做出的决定经常出错；辨不清方向或容易迷路）			
2. 缺乏兴趣、爱好了，活动减少了。比如：几乎整天和衣躺着看电视，平时讨厌外出，常闷在家里，身体懒得动，无精打采			
3. 不断重复同一件事。比如：总是提相同的问题，一句话重复多遍等			
4. 学习使用某些日常工具或家用电器（如遥控器、微波炉、VCD 等）有困难			
5. 记不清当前月份或年份			
6. 个人经济财产掌控困难（忘了如何使用存折，忘了付水、电、煤气账单等）			
7. 记不住和别人的约定。比如：和家人约好的聚会，计划去拜访亲朋好友			
8. 日常记忆和思考能力有问题。比如：自己放置的东西经常找不着，经常忘了服药，想不起熟人的名字，忘记要买的东西，忘记看过的电视、报纸、书籍的主要内容，与别人谈话时无法表达自己的意思等			
总体得分			

提示：第一栏中的“是”表示在过去的几年中在认知能力方面（记忆或者思考）出现问题。如有两个及两个以上的项目回答为“是”，很可能是记忆出了问题，建议去记忆障碍门诊就诊或向专业医师咨询。

一、AD8 使用说明

对所有回答的自发更正都是允许的，且不记录为错误。

(1) AD8 中的问题可以张贴在布告栏中用于自检，也可以由他人大声地读给受试者听，也可以在电话里询问受试者。

(2) 如果可能，AD8 最好由了解受试者的知情者来回答。但如果没有合适的知情者，AD8 也可以由患者自己回答。

(3) 当知情者回答问卷时，需要特别向他/她说明的是评价受试者的变化。

(4) 当受试者回答问卷时，需要特别向他/她说明的是评价选项相关的自身能力的改变，不需要考虑病因。

(5) 如果是念给受试者听，很重要的一点是医护人员要仔细地逐字逐句地朗读，并强调变化是基于认知障碍（而非躯体障碍）。在每单项间需要停顿 1 s 以上。

(6) 对变化发生的时间范围没有要求。

(7) 最终的分数是回答“是，有变化”的项目总数。

二、评价

AD8 不受患者年龄、教育、性别、种族的影响，不需要基线材料，可以自评也可以知情者评估，评分方法简单，耗时短，是临床医师节约时间的好帮手。

AD8 筛查本身不足以诊断痴呆。但 AD8 能非常敏感地检测出很多常见痴呆疾病的早期认知改变，包括阿尔茨海默病、血管性痴呆、路易体痴呆和额颞叶痴呆。异常范围的分数提示需要进一步的检查评估。正常范围的分数提示不太可能存在痴呆症，但不能排除是疾病的极早期。如果存在认知障碍的其他客观证据，则需要做进一步的其他检测。

Galvin(2007)连续评估 325 例就诊者，CDR=0 的占 46%，CDR=0.5 的占 31%，CDR=1 的占 15%，CDR=2 或 3 的占 8%，AD8 识别痴呆的情况见表 12-4-2。知情者评估要优于就诊者自评。Galvin 发现知情者 AD8 与认知筛查量表配合使用，可以改善认知筛查量表诊断痴呆的准确性。

表 12-4-2 痴呆知情者问卷(AD8)识别痴呆的敏感性与特异性

指 标	划界分	敏感性(%)	特异性(%)
知情者 AD8 评分	1	90	68
	2	84	93
	3	76	90
就诊者本人自评 AD8 得分	1	80	59
	2	62	73
	3	47	82

Carpenter(2011)比较了 Ottawa 3DY (O3DY,来源于加拿大健康与衰老研究的 4 个项目的筛查量表)、简易阿尔兹海默病筛查(BAS)、简短 Blessed 测验(SBT)与照料者完成的 AD8(cAD8)在老年急诊中的应用,以 MMSE<24 分作为认知受损的标准,cAD8 的识别率为 56%,cAD8 与另外 3 个筛查量表配合,并没有改善认知筛查量表的诊断准确性。

Razavi(2014)发现,区分正常组与痴呆患者,AD8 的曲线下面积为 0.953,IQCODE 为 0.930,AD8 要优于 IQCODE。以 AD8>2,IQCODE>3.4 作为划界分,AD8 有 1 例被误诊,而 IQCODE 有 27 例被误诊。同样,识别 MCI,也是 AD8 要优于 IQCODE。

北京大学精神卫生研究所李涛等(2012)完成对 AD8 的修订与验证,发现 AD8 中文版鉴别认知正常者与 AD 患者的曲线下面积为 0.93,以≥2 为认知损害的界限分值,敏感度为 93.9%,特异度为 76.0%。

目前尚缺乏 AD8 在识别 MCI 方面充分的数据支持。由于 MCI 就诊者大部分是在没有知情者陪同的情况下单独就诊,估计自评的 AD8 得分的可信度偏低。

除了 AD8,文献中亦有大量类似的量表,如痴呆初期征兆观察列表(observation list for early signs of dementia, OLD,见 Marijke HR, et al. Int J Geriatr Psychiatry. 2001, 16(4): 406 - 414.)与阿尔兹海默病问卷(Alzheimer' questionnaire, AQ,见 Malek-Ahmadi, et al. Archives of Clinical Neuropsychology. 2010,25(6): 475 - 583.)(表 12-4-3、表 12-4-4)。这些量表的目的都是帮助临床医师在患者就诊过程中早期识别 AD 患者。借助这些量表,可以更全面、更细致地注意患者的微小变化,留意其家人或照料者提供的信息。OLD 还没有评分标准与认知障碍的判断标准。AQ 的评分标准如下:0~4 分表示不用担心;5~14 分表示记忆损害,需要排除是痴呆早期表

现;大于 14 分表示可能是痴呆早期。

我们把 OLD 与 AQ 也附在本文中,有兴趣的临床医师在科研时可以比较这些不同的量表在识别 MCI 与痴呆时的效力的差异。

表 12-4-3 痴呆初期征兆观察列表(OLD)

项目
1. 总是忘记日期
2. 经常忘记短时间内的事
3. 不能重复最近听到的事或信息(如最近的检查结果等)
4. 无意识地经常重复说过的话
5. 总是在谈话中重复讲述同一件事
6. 经常想不起来特定的词语及语句
7. 说话很快就没有了中心(如话题经常转换)
8. 从回答中可以得知并未理解问题
9. 对话内容变得难以理解
10. 没有时间观念
11. 制造借口(被指出回答有错误时试图虚构)
12. 存在依赖家人的现象(向本人提问后,转向家人求助等)

表 12-4-4 阿尔兹海默病问卷(AQ)

问题	回答
1. 您的家人有记忆损害吗?	是 1□否 0□
2. 如果有,是否比前几年更严重?	是 1□否 0□
3. 是否他/她在同一天内重复同样的问题、话或故事?	是 1□否 0□
4. 您是否总是需要帮助他/她做完事或他/她总是忘记重要事情?	是 1□否 0□
5. 是否他/她每月超过 2 次会乱放东西	是 1□否 0□
6. 是否他/她找不到东西时,怀疑是别人藏起来或偷走?	是 1□否 0□
7. 是否他/她经常搞不清星期、日期、月份或一天里多次查看日期?	是 1□否 0□
8. 是否他/她在不熟悉的地方容易迷路?	是 1□否 0□
9. 是否他/她在外或旅游时容易迷糊?	是 1□否 0□
10. 是否他/她在购物找零、算账方面有问题?	是 1□否 0□
11. 是否他/她在付费、理财方面有问题?	是 1□否 0□
12. 是否他/她忘记吃药或没按要求用药?	是 1□否 0□
13. 是否他/她在驾车或骑车方面有问题或让您担忧?	是 1□否 0□
14. 是否他/她在购物找零、算账、理财方面有问题?	是 1□否 0□

(续表)

15. 是否他/她做日常的家务事有问题?	是 1□否 0□
16. 是否他/她原有的爱好(打牌、跳舞、书法等)已经不再有了?	是 1□否 0□
17. 是否他/她在熟悉的环境(小区、邻居家)中迷路?	是 1□否 0□
18. 是否他/她丧失了方向感?	是 1□否 0□
19. 是否他/她说话时找词(不是姓名)有问题?	是 1□否 0□
20. 是否他/她混淆家人或朋友的名字?	是 1□否 0□
21. 是否他/她认出熟悉的亲友有困难?	是 1□否 0□

(杨　青　郭起浩)

第十三章 日常生活能力与照料者负担量表

第一节 日常生活能力量表(ADL)

日常生活能力量表(activity of daily living scale, ADL)非常多，见表13-1-1。最简便、最常用的是由Lawton和Brody制定于1969年的版本，由躯体生活自理量表(physical self-maintenance scale, PSMS)和工具性日常生活活动量表(instrumental activities of daily living scale, IADL)组成。主要用于评定受试者的日常生活能力。ADL共有14项，包括两部分内容：一是躯体生活自理量表，共6项：上厕所、进食、穿衣、梳洗、行走和洗澡；二是工具性日常生活能力量表，共8项：打电话、购物、备餐、做家务、洗衣、使用交通工具、服药和自理经济。评分方法有很多，Lawton和Brody 1969年的版本中(表13-1-2)，1分为完全独立完成，0分为部分完成或完全依赖，如果要增加分级，根据不同权重区分，如“使用交通工具”分5级而“服药”分3级。在Lawton 1982年发表的9项MAI版本，每个项目都是1～3分。评定时按表格逐项询问，如果受试者因故不能回答或不能正确回答(如痴呆或失语)，则可根据家属、护理人员等知情人的观察评定。如果无从了解，或从未做过的项目，例如没有电话也从来不打电话，记(9)，以后按研究规定处理。评定结果可按总分、分量表分和单项分进行分析。ADL受多种因素影响，年龄、视觉、听觉或运动功能障碍、躯体疾病、情绪低落等，均影响日常生活功能。对ADL结果的解释应谨慎。

表 13-1-1 日常生活能力量表(ADL)

简称	全称	编制者
ADL-PI	日常生活能力-预防工具	Galasko,2006
ADLQ	日常生活能力问卷	Johnson,2004
ADCS-ADL19/23	ADCS(AD 协作中心)编制的 19 项或 23 项 ADL 清单	Galasko,1997
B-ADL	Bayer 日常生活能力	Hindmarch,1998
Bristol ADL	Bristol 日常生活能力	Bucks,1996
BANS-S	Bedfords AD 护理严重度量表	Bellelli,1997
Barthel ADL	Barthel 日常生活能力量表	Mahoney&Barthel,1965
DAD	痴呆残疾评估	Gelinas,1999
GERRI	基于家属的老年活动能力评估量表	Schwartz,1983
IADL	工具性日常生活能力量表	Lawton&Brody,1969
IDDD	痴呆患者日常生活能力变化晤谈量表	Teunisse,1991
Katz ADL	Katz 日常生活能力独立指数	Katz,1963
PGDRS	老年人心理依赖评估量表	Wilkinson&Graham-White,1980
RDRS	快速残疾评估量表	Linn,1976
PSMS	躯体自我维持量表	Lawton&Brody,1969
SMAF	功能自主评估系统	Hebert,2001
Weintraub ADL	Weintraub 日常生活能力量表	Weintraub,1986

表 13-1-2 日常生活能力量表(Lawton 和 Brody,1969 年)

评分	工具性日常生活能力量表(IADL)评分项目
□	A. 使用电话的能力
1	1. 自己主动操作电话——查号码、拨号等
1	2. 能拨几个熟悉的号码
1	3. 能接电话但不拨电话
0	4. 完全不会使用电话
□	B. 购物
1	1. 独立处理所有购物需要
0	2. 独立进行少量购物
0	3. 任何购买途中都需要陪伴
0	4. 完全不能购物
□	C. 做饭
1	1. 独立计划、准备并做出适量的饭
0	2. 如果供给原料能准备足量饭
0	3. 加热、服务或做饭,或做饭但不能保持适量

（续表）

评分	工具性日常生活能力量表(IADL) 评分项目
0	4. 需要把饭准备好并做好
□	D. 主持家务
1	1. 独立主持家务或偶尔需要帮助(如重活需要家人帮助)
1	2. 做日常轻体力家务如洗碗、铺床
1	3. 做日常轻体力家务但不能保持可接受的清洁水平
1	4. 所有家务都需要帮助
0	5. 不参与任何家务
□	E. 洗涤
1	1. 能完成个人洗澡
1	2. 洗小件衣物——洗袜子等
0	3. 所有洗涤必须靠其他人完成
□	F. 交通方式
1	1. 独立乘坐公共车辆或驾驶小汽车
1	2. 乘出租车旅行，但不用公共汽车
1	3. 有他人陪伴时乘坐公共汽车
0	4. 在他人帮助下，只能乘出租车或汽车旅行
0	5. 完全不能旅行
□	G. 对自己医疗的责任心
1	1. 能认真按照正确时间、剂量吃药
0	2. 如果预先准备了每次剂量能吃药
0	3. 不能准备自己吃的药
□	H. 理财能力
1	1. 独立处理财务(预算，写支票，付租金、账单，去银行)，收集和保持收入渠道
1	2. 管理日常购物，但在银行事务和大的购物等情况下需要帮助
0	3. 不能处理财务

评分	躯体性自理能力量表(PSMS)评分项目
□	A. 大小便卫生
1	1. 在盥洗室能完全自理，无失禁
0	2. 在自我清洁方面需要提醒，或需要帮助，或有少量的事故发生(至少一周1次)
0	3. 睡眠时弄脏或弄湿，超过一周1次
0	4. 清醒时弄脏或弄湿，超过一周1次
0	5. 大小便失禁
□	B. 吃饭
1	1. 吃东西不需帮助
0	2. 吃东西需要少量帮助和(或)需要准备特殊食物，或在餐后清洁时需要帮助
0	3. 自己吃饭需要中等帮助，并且不整齐

（续表）

评分	工具性日常生活能力量表(IADL) 评分项目
0	4. 所有的就餐需要多方面的帮助
0	5. 完全不能自己吃饭，并抵抗他人喂食
□	C. 穿衣
1	1. 穿衣服，脱衣服，并能从自己的衣柜选择衣服
0	2. 自己穿衣服和脱衣服，需要少量帮助
0	3. 在穿衣服和选择衣服方面需要中等帮助
0	4. 在穿衣服上需要多帮助，但配合他人的帮助
0	5. 完全不能自己穿衣服，并对他人的帮助有抵触
□	D. 梳理　（整洁，头发，指甲，手，脸，衣服）
1	1. 总是穿戴整洁，妆饰恰当，无需帮助
0	2. 能自己恰当装饰，偶尔需要少量帮助，如修胡须
0	3. 在梳理上需要中等合理的帮助或指导
0	4. 所有梳理事务都需要帮助，但在他人帮助后能保持整洁
0	5. 主动反对他人所有的帮助梳理的努力
□	E. 躯体步行
1	1. 步行到场地或市区
0	2. 在居住区内步行，或在一条街道附近步行
0	3. 步行需要帮助(选择1个)a(　)手杖，d(　)步行者，e(　)轮椅 1——出入无须帮助 2——出入需要帮助
0	4. 不用支持坐在椅子上或轮椅上，但没有帮助就不能自己推进
0	5. 一半多的时间卧床不起
□	F. 洗澡
1	1. 自己洗澡(盆浴，淋浴，擦浴)，无须帮助
0	2. 在进出浴盆时需要帮助
0	3. 只洗脸和手，不能洗身体其他部位
0	4. 不能自己洗澡，但配合他人给他洗澡
0	5. 不能自己洗澡，并且抵制让他保持清洁的努力

除了 Lawton 和 Brody 版的 ADL，本节介绍张明园教授修订的 ADL 中文版(表 13-1-3)，Bristol 日常生活能力量表(表 13-1-4、表 13-1-5)和 ADCS-ADL(表 13-1-6)，这些量表的共同特点是针对老人的躯体功能(如穿衣、梳头、上厕所、走楼梯和剪指甲等)和使用工具能力(如打电话、做饭菜和购物等)，而门诊相当多的就诊者只是工作或职业、社交活动能力的下降，按照 DSM-Ⅳ-R 的痴呆诊断标准，这些复杂能力下降(而不是简单的

日常活动能力减退）才是痴呆的早期诊断依据之一，但是，每个就诊者的职业千差万别，要对职业能力进行量化评估是很困难的，所以，笔者设计了一个自评的简易的复杂社会功能量表(PFS)（表 13-1-7），其信度与效度正在验证中。

表 13-1-3　日常生活活动量表(ADL，张明园修订版)

项　　目	评分	项　　目	评分
1. 自己搭公共车辆	1 2 3 4	11. 上下楼梯	1 2 3 4
2. 到家附近的地方去(步行范围)	1 2 3 4	12. 上下床，坐下或站起	1 2 3 4
3. 自己做饭(包括生火)	1 2 3 4	13. 提水煮饭、洗澡	1 2 3 4
4. 做家务	1 2 3 4	14. 洗澡(水已放好)	1 2 3 4
5. 吃药	1 2 3 4	15. 剪脚趾甲	1 2 3 4
6. 吃饭	1 2 3 4	16. 逛街、购物	1 2 3 4
7. 穿衣服、脱衣服	1 2 3 4	17. 定时去厕所	1 2 3 4
8. 梳头、刷牙等	1 2 3 4	18. 打电话	1 2 3 4
9. 洗自己的衣服	1 2 3 4	19. 处理自己钱财	1 2 3 4
10. 在平坦的室内走	1 2 3 4	20. 独自在家	1 2 3 4

注：提问有关患者平常每天需要做的事件。评分标准：1 分为自己可以做；2 分为有些困难；3 分为需要帮助；4 分为根本没法做。

表 13-1-4　Bristol 日常生活能力量表

项　目	0	1	2	3	4	项　目	0	1	2	3	4
1. 做饭菜						11. 行走					
2. 吃饭						12. 时间定向					
3. 准备饮料						13. 空间定向					
4. 喝饮料						14. 对话交流					
5. 穿衣						15. 打电话					
6. 卫生						16. 做家务					
7. 刷牙						17. 购物					
8. 洗澡						18. 理财					
9. 上厕所						19. 业余爱好					
10. 移动						20. 交通工具					

注：这份问卷的设计目的是为了揭示有一种或其他形式记忆障碍的人的日常生活能力。对于每一项行为(第 1～20 项)，0～3 选项对应不同水平的能力。请回想一下最近 2 周的情况，圈出代表您的亲戚或朋友相应行为能力的数字。每一项行为只能圈一个数字(如果对圈哪个数字有疑问的话，选择能代表最近 2 周平均情况的能力水平)。

表 13-1-5　Bristol 日常生活能力量表项目评分说明

项目	说明
1. 做饭菜	0 分＝能根据需要选择和准备食物；1 分＝提供配料的情况下能做饭菜；2 分＝一步一步提示的情况下能做饭菜；3 分＝即使在提示和指导的情况下也不能做饭；4 分＝不符合
2. 吃饭	0 分＝能使用正确的餐具正常地吃饭；1 分＝在食物容易食用和(或)使用汤匙的情况下能正常地吃饭；2 分＝用手抓饭吃；3 分＝需要喂食；4 分＝不符合
3. 准备饮料	0 分＝能根据需要选择及准备饮料；1 分＝有现成配料的情况下能准备饮料；2 分＝一步一步提示的情况下能准备饮料；3 分＝即使在提示和指导的情况下也不能制作饮料；4 分＝不符合
4. 喝饮料	0 分＝能正常地自己喝饮料；1 分＝提供帮助、杯子或吸管的情况下能正常地喝饮料；2 分＝即使在提供帮助的情况下也不能正常地喝饮料，但努力尝试喝；3 分＝不能自己喝，必须由看护喂；4 分＝不符合
5. 穿衣	0 分＝能选择合适的衣服并且能自己穿；1 分＝穿衣的顺序错误和(或)穿错前后面和(或)穿脏衣服；2 分＝不能自己穿衣，但能配合地移动肢体；3 分＝不能配合，完全需要别人帮忙穿衣；4 分＝不符合
6. 卫生	0 分＝能定期自己清洗；1 分＝提供肥皂、面巾、毛巾等的情况下能自己清洗；2 分＝在提示和指导的情况下能自己清洗；3 分＝不能自己清洗，完全需要帮助；4 分＝不符合
7. 刷牙	0 分＝能定期自己刷牙或假牙；1 分＝提供恰当提示的情况下能自己刷牙或假牙；2 分＝需要一些帮助，把牙膏挤在牙刷上，把牙刷放到牙齿上等；3 分＝完全需要帮助；4 分＝不符合
8. 洗澡	0 分＝能定期自己洗澡；1 分＝需要帮忙打开淋浴开关，但能自己洗澡；2 分＝需要监督和指导才能洗澡；3 分＝完全需要帮助；4 分＝不符合
9. 上厕所	0 分＝需要时能自己上厕所；1 分＝需要带去厕所并提供帮助；2 分＝小便或大便失禁；3 分＝小便和大便失禁；4 分＝不符合
10. 移动	0 分＝能自己坐到椅子上或从椅子上站起来；1 分＝能自己坐到椅子上，但站起来需要帮助；2 分＝坐到椅子上和站起来都需要帮助；3 分＝坐到椅子上和站起来完全依赖于他人的帮助；4 分＝不符合
11. 行走	0 分＝能自己行走；1 分＝行走时需要辅助物(如需要家具、手臂支撑)；2 分＝使用辅助器移动(如支架、拐杖等)；3 分＝不能行走；4 分＝不符合
12. 时间定向力	0 分＝时间、昼夜、日期等定向力完整；1 分＝不知道时间、昼夜、日期等，但似乎并不关心；2 分＝重复地问时间、昼夜、日期；3 分＝分不清白天和晚上；4 分＝不符合
13. 空间定向力	0 分＝对环境的定向力完整；1 分＝只对熟悉的环境具有定向力；2 分＝在家里迷路，需要提醒浴室的位置等；3 分＝不认识自己的家并试图离开；4 分＝不符合
14. 交流	0 分＝能正常地对话；1 分＝能理解他人的话并试图配合手势用言语回答；2 分＝能让他人理解自己的意思，但难以理解他人的话；3 分＝不回答他人或不与他人交流；4 分＝不符合
15. 打电话	0 分＝正常地使用电话，包括正确地拨号；1 分＝口头告知号码或写下号码或事先拨好号码的情况下能使用电话；2 分＝能接电话，但不能打电话；3 分＝完全不能或不愿使用电话；4 分＝不符合

（续表）

项目	说明
16. 做家务/栽培花木(园艺)	0分=能像以前一样做家务或栽培花木;1分=能做家务或栽培花木,但不如以前;2分=即使提供很多指导,参与时仍然受到限制;3分=不愿或不能参与以前的活动;4分=不符合
17. 购物	0分=能像以前一样购物;1分=有或没有清单的情况下只能够买一二样东西;2分=不能自己购物,有人陪伴的情况下愿意参与;3分=即使有人陪伴也不愿参与购物;4分=不符合
18. 理财	0分=能像以前一样可靠地处理自己的钱财;1分=不能写支票,但能签名并能分辨金额大小/币值;2分=能签名,但不能分辨金额大小/币值;3分=不能签名、分辨金额大小/币值;4分=不符合
19. 游戏或业余爱好	0分=像以前一样参加娱乐活动;1分=能参加娱乐活动,但需要指导或监督;2分=勉强能参加娱乐活动,很迟钝,需要哄;3分=不能或不愿再参加娱乐活动;4分=不符合
20. 交通工具	0分=能自己开车、骑自行车或使用公共交通工具;1分=不能开车,但能使用公共交通工具或骑自行车等;2分=不能单独使用公共交通工具;3分=即使有人陪伴也不能或不愿使用交通工具;4分=不符合

表13-1-6 ADCS-ADL(AD协作中心Galasko等1997编制)

序号	项目	请选择	评分
1	关于吃东西,哪一项最恰当地描述了他/她的通常表现	3 吃东西时不需要帮助,并且没有困难 2 吃东西时不需要帮助,但有些困难 1 吃东西时不需要帮助,但很困难 0 患者通常或总是由别人喂食	□
2	关于行走(或坐在轮椅上四处移动),哪一项最恰当地描述了他/她的通常表现	3 不需要身体上的帮助就能在户外四处活动 2 不需要身体上的帮助就能在房间里四处活动 1 不需要帮助就能从床上转移到椅子上 0 行走或转移时需要身体上的帮助	□
3	关于如厕时的大小便功能,哪一项最恰当地描述了他/她的通常表现	3 不需要监护或帮助就能做一切必要的事情 2 需要监护,但不需要身体上的帮助 1 需要身体上的帮助,通常能控制大小便 0 需要身体上的帮助,通常不能控制大小便	□
4	关于洗澡,哪一项最恰当地描述了他/她的通常表现	3 不需要提醒或身体上的帮助就能洗澡 2 不需要身体上的帮助,但是需要监护/提醒才能完成洗澡的全部过程 1 需要少许身体上的帮助(如:在洗头发时)才能完成洗澡的全部过程 0 完全需要由别人帮助	□
5	关于个人卫生,哪一项最恰当地描述了他/她的最佳表现	3 不需要身体上的帮助就能清洁和修剪手指甲 2 不需要身体上的帮助就能梳头发 1 不需要身体上的帮助就能洗脸和洗手 0 需要有人帮忙梳头发、洗脸、洗手以及修剪手指甲	□

（续表）

序号	项 目	请 选 择	评分
6	关于穿衣		
6A	患者挑选过白天要穿的第一套衣服吗？	是 否 不知道 0 0	□
	如果是，请选择	3 不需要监护或帮助 2 需要监护 1 需要帮助	
6B	哪一项最恰当地描述了他/她的通常表现	4 不需要监护或身体上的帮助就能完成穿衣的全部过程 3 需要监护才能完成穿衣的全部过程，但不需要帮助 2 只在扣纽扣、扣钩或拴鞋带时需要帮助 1 如果衣服不用系扣，那么不需要帮助就能穿衣 0 不管穿什么类型的衣服总是需要帮助	□
7	患者用过电话吗？	是 否 不知道 0 0	□
	如果是，哪一项最恰当地描述了他/她的最高表现水平	5 能在查询电话号码簿上的号码后打电话；或者能在拨打查号服务号码后打电话 4 在不看电话号码簿或电话号码单的情况下，只能拨打熟悉的号码 3 在使用电话号码簿或电话号码单的情况下，只能拨打很熟悉的号码 2 能接电话；不能打电话 1 不能接电话，但别人把电话给他/她时能说话	
8	患者看过电视吗？ 如果是，就问下列所有问题	是 否 不知道 0 0	□
		a）患者通常会选择或要求看不同的节目或他/她最喜欢的表演吗？ 1 0 0	□
		b）患者通常会在看节目时谈论节目的内容吗？ 1 0 0	□
		c）患者会在看完节目后的1天(24 h)之内谈论节目的内容吗？ 1 0 0	□
9	患者曾参与对话或闲谈至少5 min吗？注意：患者不需要发起谈话。	是 否 不知道 0 0	□
	如果是，哪一项最恰当地描述了通常他/她参与谈话的程度	3 通常说与话题相关的事情 2 通常说与话题无关的事情 1 很少或从不说话	
10	患者在饭后或吃小吃后收拾过桌子吗？	是 否 不知道 0 0	□
	如果是，哪一项最恰当地描述了他/她的通常表现	3 不需要监护或帮助 2 需要监护 1 需要帮助	

（续表）

序号	项目	请选择				评分
11	患者通常能在家里找到他/她的个人物品吗？		是	否 0	不知道 0	□
	如果是，哪一项最恰当地描述了他/她的通常表现	3　不需要监护或帮助 2　需要监护 1　需要帮助				
12	患者为自己获取过热饮料或冷饮料吗？（冷饮料包括水）		是	否 0	不知道 0	□
	如果是，哪一项最恰当地描述了他/她的最高表现水平	3　通常不需要身体上的帮助就能制作热饮料 2　如果有人热了水，通常能制作热饮料 1　通常不需要帮助就能获取冷饮料				
13	患者在家里为自己做过饭或小吃吗？		是	否 0	不知道 0	□
	如果是，哪一项最恰当地描述了他/她准备食物时的最高水平	4　需要很少的帮助或不需要帮助就能烹煮食物或用微波炉热食物 3　需要大量帮助才能烹煮食物或用微波炉热食物 2　没有烹煮食物或用微波炉热食物，将一些食物混合或组合在一起做成一顿饭或小吃（比如做三明治） 1　没有混合或烹煮食物，自己能获取食物				
14	患者把垃圾或杂物扔在家里恰当的地方或装垃圾的容器里了吗？		是	否 0	不知道 0	□
	如果是，哪一项最恰当地描述了他/她的通常表现	3　不需要监护或帮助 2　需要监护 1　需要帮助				
15	患者从他/她家里出来四处活动（或出行）吗？		是	否 0	不知道 0	□
	如果是，哪一项最恰当地描述了他/她的最佳表现	4　独自出门，离家至少 1 km 远 3　独自出门，但在离家 1 km 的范围内 2　不管离家多远，只在有人陪伴和监护时才出门 1　不管离家多远，只在获得身体上的帮助的情况下才出门				
16	患者曾购过物吗？ 如果是，就问 A 和 B：		是	否 0	不知道 0	□
	A）哪一项最恰当地描述了患者通常选择物品的方式	3　不需要监护或身体上的帮助 2　需要一些监护或身体上的帮助 1　根本就不挑选，或者主要挑选的是胡乱拿的物品或不恰当的物品				□
	B）患者通常在不需要监护或身体上的帮助的情况下就能付钱购买物品吗？		是	否 0	不知道 0	□

（续表）

序号	项　目	请　选　择	是	否	不知道	评分
17	患者能想起和其他人（比如亲戚、朋友、医生或理发师等）的预约或见面吗？		是	否 0	不知道 0	□
	如果是，哪一项最恰当地描述了他/她事先想起见面的记忆能力	3　通常都能记得，可能需要用书写的方式提醒，比如写在便条上、日记里或者日历上 2　只能在当天口头提醒之后还记得预约的事情 1　尽管当天口头提醒过，但通常还是忘记了				
注意：如果患者住在养老院就不问问题 18，跳到问题 19；并在此处打钩　□						
18	患者曾被独自留下吗？（家里有人照料不等同于患者被独自留下） 如果是，就问下列所有问题		是	否 0	不知道 0	□
		a）患者白天曾被独自留在家外 15 min 或更长的时间吗？	1	0	0	□
		b）患者白天曾被独自留在家里 1 h 或更长的时间吗？	1	0	0	□
		c）患者白天曾被独自留在家里少于 1 h 吗？	1	0	0	□
19	患者谈论过最近的事情吗？（即发生在过去 1 个月里或过去 4 周里的事情或事件。） 如果是，就问下列所有问题		是	否 0	不知道 0	□
		a）患者谈论的是他/她听过、读过或在电视上看过的，但没有参与过的事情吗？	1	0	0	□
		b）患者谈论的是他/她在户外与家人、朋友或邻居一起参与的事情吗？	1	0	0	□
		c）患者谈论的是他/她在家里参与或看到的事情吗？	1	0	0	□
20	患者读杂志、报纸或书籍时一次会超过 5 min 吗？ 如果是，就问下列所有问题		是	否 0	不知道 0	□
		a）患者通常在阅读时或阅读之后很快（少于 1 h）就会谈论他/她读到的细节吗？	1	0	0	□
		b）患者通常在阅读完 1 h 或更长的时间后，会谈论他/她读到的东西吗？	1	0	0	□
21	患者曾写过东西吗？ 注意：如果患者仅在受到鼓励或获得帮助后才写东西，回答仍应该为“是”		是	否 0	不知道 0	□
	如果是，哪一项最恰当地描述了他/她写过的最复杂的东西	3　其他人能看懂的信件或长便条 2　其他人能看懂的短便条或留言 1　他/她的签名或名字				

（续表）

序号	项　目	请　选　择	评分
22	患者进行过消遣、从事过爱好或玩过游戏吗？	是　否　不知道 0　0	□
	如果是，哪些是他/她进行过的消遣 注意：在本量表里，走路不算作爱好/消遣 询问所有内容，将所有的符合项打钩	□ a）棋牌游戏（包括麻将、扑克、象棋） □ b）太极拳 □ c）乐器 □ d）阅读 □ e）游泳 □ f）猜谜游戏 □ g）针织 □ h）园艺（如种花、浇花等） □ i）家庭修理 □ j）艺术（包括绘画、书法、舞蹈等） □ k）缝纫 □ l）门球 □ m）钓鱼 □ n）其他______	
	如果是，通常患者如何进行最常做的消遣	3　不需要监护或帮助 2　需要监护 1　需要帮助	□
22A	如果患者只在有人照料的时候才从事爱好/进行消遣，请勾选		□
23	患者用过家用电器做家务吗？	是　否　不知道 0　0	□
	如果是，询问以下所有各项，勾选使用过的家用电器	□ a）洗衣机 □ b）熨斗 □ c）燃气灶 □ d）电蚊拍 □ e）饮水机 □ f）微波炉 □ g）吸尘器 □ h）电饭锅 □ i）食品加工器 □ j）其他______	
	如果是，对于最常用的电器，哪一项最恰当地描述了患者通常如何使用它们	4　不需要帮助，如果需要进行操作，操作不仅仅只是开、关控制 3　不需要帮助，但是只操作开、关控制 2　需要监护，但不需要帮助 1　需要帮助	□
24		总分（0～78）	
25		“不知道”回答的次数	

注：请知情者回答在过去4周里患者的表现。

表 13-1-7　复杂社会功能量表(PFS)

功　能	下降幅度(%)	可　能　原　因			
		躯体疾病	记忆减退	脑力迟钝	情绪异常
1. 炒股					
2. 理财(除外炒股)					
3. 写作(如做笔记)					
4. 讲课或讲座					
5. 打麻将					
6. 打扑克					
7. 下象棋					
8. 绘画与书法					
9. 演奏乐器					
10. 修理电器故障					
总体判断					

注：复杂社会功能量表(PFS)是笔者编制的自评问卷，请患者根据近 3 个月的情况回答。共列举了 10 种需要脑力劳动的复杂功能(如果你从来不参与或从事该项活动，请空着)。在能力下降原因里，脑力迟钝指思维迟钝、延缓或阻滞的感觉；情绪异常指焦虑、抑郁、失眠等。

(郭起浩)

第二节　照料者负担量表

在世界范围内，家庭系统一直以来是痴呆照料的基石。国内外研究发现痴呆家属照料者较非痴呆家属照料者经历更多的应激，这些应激包括客观应激源(stressors)和主观角色负担(strain)。客观应激源包括照料者的人口统计学特征、被照料者的认知、行为和躯体功能、照料者角色对家庭和职业的冲突、经济上的压力，以及因照料而导致的社交孤立。主观角色负担包括照料过程的情绪体验、关系剥夺、照料的主观评价、动机和影响。上述这些应激引发家属照料者明显的躯体和心理疾患。澄清和评估这些应激是制定对痴呆患者及其家属照料者有效的服务和干预项目中不可或缺的部分。常用的评估照料者应激工具中，Zarit 照料者负担量表(Zarit caregiver burden interview，ZBI)、照料者负担问卷(caregiver burden inventory，CBI)和照料压力问卷(caregiving stress inventory，CStI)的研发过程均来自对痴呆患者家属照料者的研究，因而被广泛地应用于该人群的相关研究，包括评价照料负担、主要

影响因子、对照料者健康相关结局的影响等。由于目前上述3个工具中，只有Zarit照料者负担量表(以下简称ZBI)已有中国内地学者修订的中文版，因此本节主要介绍ZBI及该中文版信效度研究结果。

ZBI由美国学者Zarit(1980)根据照料负担测量的理论研制而成。ZBI作为在痴呆患者家属照料者的研究中最常用的工具，主要用于评价照顾者感受到的照料负担及其影响因素包括照料者经常提到的影响其健康、经济、社交、与被照料者的关系的问题。ZBI最初的版本共有29个条目，后经修订发展为22项、18项、12项共3个版本。目前广泛使用的为22个条目的版本(表13-2-1)。12条目的版本主要作为筛查工具而使用。22条目版的ZBI包括角色负担(role strain)和个人负担(personal strain)2个维度共22个条目。角色负担由条目1、4、5、8、9、14、16、17、18、19、20、21组成，评估照料角色的需求，而个人负担(由条目2、3、6、11、12、13组成)评估的是照料者本人对照料角色的自我肯定。量表评分方法为将照顾者感受到的照料负担分为下述5个等级：0＝从来没有过，1＝很少(每月几次)，2＝有时(每周1～3次)，3＝相当频繁(每周4～5次)和4＝几乎总是(几乎每天都有这种感觉)。角色负担和个人负担的总分分别在0～48分和0～36分，分数越高提示照料者负担越重。大量研究发现ZBI总分与照料者身心健康、活力和社交功能呈显著负相关。ZBI还被用于筛查照料者抑郁症风险和评价是否需对其进行干预的研究。Schreiner等人采用ZBI、GDS和CES-D对卒中(n=80)、慢阻肺(n=48)和其他功能残疾(70例，其中70%为痴呆患者)共3组患者家属照料者进行研究，以确定预测家属照料者抑郁症风险的ZBI临界值(cut-off)。结果发现预测抑郁症风险的ZBI临界值为24～26，ZBI总分24时，72%照料者有抑郁症可能。

表13-2-1 Zarit照料者负担量表(ZBI)

下列描述反映了人们在照顾其他人时的一些感受。请根据您自己的感受进行选择。请在相应的数字上打“√”

0＝从来没有过；1＝很少(每月几次)；2＝有时(每周1～3次)；
3＝相当频繁(每周4～5次)；4＝几乎总是(几乎每天都有这种感觉)

您觉得…	感受的频率				
	0	1	2	3	4
1. 老人要求的帮助比他/她实际所需要的多					
2. 因为把时间花在照顾老人身上，自己的时间不够用					
3. 既要照顾老人又要承担其他家庭和工作责任使您紧张					

（续表）

您觉得…	感受的频率				
	0	1	2	3	4
4. 老人的一些行为让人难堪					
5. 当老人在您身边时，您感到生气					
6. 老人对您与其他家庭成员的关系产生了负面的影响					
7. 您为老人的前途担忧					
8. 老人很依赖您					
9. 当老人在您身边时，您觉得紧张					
10. 因为照顾老人，您感到自己的健康受损了					
11. 因为照顾老人，您的个人生活不如期望的好了					
12. 因为照顾老人，您的社交生活受到了影响					
13. 您不太愿意叫朋友过来玩了					
14. 老人在期待您照顾他/她，就好像您是他/她唯一的依靠					
15. 您的收入不够照顾老人					
16. 您没有能力再照顾老人很长时间了					
17. 您的生活被搅乱了					
18. 您希望把照顾老人的任务交给别人					
19. 您不知道拿他/她怎么办					
20. 您应该为他/她做得更多					
21. 您可以在照顾他/她的方面做得更好					
22. 总的来说，您觉得照顾老人对您是一个负担					

中文版 ZBI(表 13-2-2)由王烈等人修订，信效度分析显示量表总的和两维度 Cronbach's α 系数均大于 0.7(0.75～0.88)；条目敏感性分析结果提示量表反应度较好(Cronbach's α 系数为 0.87～0.89)；各条目分与总分相关分析结果发现条目 4 (0.230)、6(0.249)、20(0.249)和 21(0.160)的相关系数低于 0.3，提示上述条目内容与量表内容相关性较低；各维度和条目间 Cronbach's α 系数中仅条目 5 与内部条目(角色负担)的 α 系数(0.273)小于外部条目(个人负担)的 α 系数(0.361)。上述信效度研究结果提示中文版 ZBI 信效度较好，但笔者认为还需对该中文版进一步修订。

表 13-2-2　照料者应激过程评估量表

下列描述反映了人们在照顾其他人时的一些感受。请根据您自己的感受进行选择。请在相应的数字上打“√”

0=从来没有过；1=很少(每月几次)；2=有时(每周 1～3 次)；
3=相当频繁(每周 4～5 次)；4=几乎总是(几乎每天都有这种感觉)

您觉得	感受的频率				
	0	1	2	3	4
1. 老人要求的帮助比他/她实际所需要的多					
2. 因为把时间花在照顾老人身上，自己的时间不够用					
3. 既要照顾老人又要承担其他家庭和工作责任使您紧张					
4. 老人的一些行为让人难堪					
5. 当老人在您身边时，您感到生气					
6. 老人对您与其他家庭成员的关系产生了负面的影响					
7. 您为老人的前途担忧					
8. 老人很依赖您					
9. 当老人在您身边时，您觉得紧张					
10. 因为照顾老人，您感到自己的健康受损了					
11. 因为照顾老人，您的个人生活不如期望的好了					
12. 因为照顾老人，您的社交生活受到了影响					
13. 您不太愿意叫朋友过来玩了					
14. 老人在期待您照顾他/她，就好像您是他/她唯一的依靠					
15. 您的收入不够照顾老人					
16. 您没有能力再照顾老人很长时间了					
17. 您的生活被搅乱了					
18. 您希望把照顾老人的任务交给别人					
19. 您不知道拿他/她怎么办					
20. 您应该为他/她做得更多					
21. 您可以在照顾他/她的方面做得更好					
22. 总的来说，您觉得照顾老人对您是一个负担					

此外，对照料者负担的定义和影响因素的研究发现，ZBI 对照料者负担的评估着重于照顾者个人及其家庭范畴，是单维的。对照料者负担的评估内容应包含更多的维度，比如常见的照料任务、社交限制情况、对照料者本人身心损害、照料者对照料经历引起的应激和负担的主观感受。与影响因素相关的评估范畴还需要包含社区及更大的范围。

除了 ZBI，笔者翻译修订了 Pearlin(1990)编制的“照料者应激过程评估(measures specifying the caregiver stress process)”(表 13-2-3)。前 3 个维度(角色负担、角色能力和社交受损)被称为“原发性应激”，后 3 个维度(家庭冲突、工作压力和经济压力)被称为“继发性应激”。最后的“观念”部分是笔者加上的。这个量表的优点是建立在全面的应激加工理论模型基础上的，是多维评估。其缺点是，把照料完全看成是负性事件，没有看到照料带来的正面影响。中文版的信度和效度正在验证中。

表 13-2-3 照料者应激过程评估量表

(一) 有关照料患者，你的想法是				
1. 希望患者对你的生活没有影响	一直是	有时是	偶尔是	从不这么想
2. 患者对你是一个困境	一直是	有时是	偶尔是	从不这么想
3. 你希望你刚好离家了	一直是	有时是	偶尔是	从不这么想
(二) 有关照料患者，以下的说法				
1. 晚上上床时，你已经筋疲力尽	非常同意	基本同意	有时是	非常不同意
2. 你本来可以做更多事情	非常同意	基本同意	有时是	非常不同意
3. 你留给自己的时间没有了	非常同意	基本同意	有时是	非常不同意
4. 你努力照料患者，但是感到这是徒劳	非常同意	基本同意	有时是	非常不同意
(三) 你在个人关系上的损失				
1. 你丧失了与患者的相互信任	非常同意	基本同意	有时是	非常不同意
2. 患者像换了一个人	非常同意	基本同意	有时是	非常不同意
3. 患者不再了解你、体谅你	非常同意	基本同意	有时是	非常不同意
(四) 对以下问题，你家庭中每个成员的看法是否一致				
1. 患者记忆障碍的严重度	非常一致	有时不一致	相当不一致	非常不一致
2. 是否需要当心患者的安全	非常一致	有时不一致	相当不一致	非常不一致
3. 患者的生活自理能力	非常一致	有时不一致	相当不一致	非常不一致
4. 是否到医院就诊	非常一致	有时不一致	相当不一致	非常不一致
(五) 除了你，其他家庭成员				
1. 没有花时间在患者身上	一直不是这样	偶尔这样	有时这样	一直这么做
2. 没有分担你对患者的照料	一直不是这样	偶尔这样	有时这样	一直这么做
3. 不重视患者	一直不是这样	偶尔这样	有时这样	一直这么做
4. 对患者缺乏耐心	一直不是这样	偶尔这样	有时这样	一直这么做
5. 对患者不闻不问	一直不是这样	偶尔这样	有时这样	一直这么做
6. 不提供任何帮助	一直不是这样	偶尔这样	有时这样	一直这么做
7. 对你的照料工作不满意	一直不是这样	偶尔这样	有时这样	一直这么做
8. 给你多余的或讨厌的建议	一直不是这样	偶尔这样	有时这样	一直这么做

（续表）

（六）对于你目前的工作状况，以下说法你是				
1. 你工作的精力减退了	非常同意	基本同意	不同意	非常不同意
2. 你的工作出现了更多失误	非常同意	基本同意	不同意	非常不同意
3. 你对自己的工作质量不满意	非常同意	基本同意	不同意	非常不同意
4. 你在工作时，为患者焦虑担心	非常同意	基本同意	不同意	非常不同意
5. 在工作时间与患者通电话	非常同意	基本同意	不同意	非常不同意
（七）与患者得病前相比，你的经济负担				
1. 患者的收入是	明显减少	有减少	没有变化	有增加
2. 你的收入是	明显减少	有减少	没有变化	有增加
3. 你的家庭总收入是	明显减少	有减少	没有变化	有增加
（八）你的观念				
1. 患痴呆是一种耻辱，应该保密	非常同意	基本同意	不同意	非常不同意
2. 痴呆是一种遗传病	非常同意	基本同意	不同意	非常不同意
3. 痴呆就是日常生活不能自理	非常同意	基本同意	不同意	非常不同意

（张曙映　郭起浩）

第三节　照料者自我效能问卷(SEQ)

中国预计将成为阿尔茨海默病患者绝对数量最多的国家，研究表明，在中国90%以上的痴呆患者将由其家属照料。国内外文献都表明痴呆患者的认知和躯体损害，特别是患者病程中高达90%的行为和精神症状(BPSD)是损害家属照料者身心健康的最重要因素。在做好对痴呆患者早发现、早诊断和早治疗的同时，通过支持措施对影响其家属照料者身心健康的风险因素进行早预防、早发现，提高家属照料者尤其是配偶的健康自我管理能力，使痴呆患者得到良好的居家照料，对改善这两类群体生活质量具有重要意义。

Parker等通过对相关干预结果指标的研究发现，抑郁症状、主观健康、自我效能和照料负担可能是评判干预有效性的指标。较其他结果指标，照料者自我效能(self-efficacy)还是一个较新的评价变量。自我效能的概念来源于Bandura(班杜拉)的社会学习认知理论。自我效能强调个体对自身在特定情境下执行某一特定行为的能力的判断。这种判断(即自我效能)决定了个体行为的动机和坚持性。近十几年来，国内外研究发现照料者自我效能对照料者

身心健康具有保护作用。相关干预研究发现，提高照料者自我效能，有助于增强个体对环境的调整和适应能力，改善照料者的健康结果。本节介绍痴呆照料者自我效能与患者各项损害、照料者的社会支持和健康相关结果间关系，以及自我效能评价工具的研究进展。

一、自我效能与痴呆损害、社会支持的关系

在照料痴呆患者的过程中，患者的认知损害、躯体功能和行为障碍，以及照料者的社会支持是影响照料者自我效能感的主要情景因素。多数研究关注患者行为障碍对照料者自我效能感的影响，患者其他损害如认知功能、日常生活能力（ADLs）和工具性日常生活能力（IADLs）对自我效能影响的报道相对有限。Pinquart 的 meta 分析发现患者的行为障碍是减弱照料者自我效能感的最重要因素。由于采用不同自我效能评估工具，患者各项损害与特定的自我效能间关系的研究结果也不一致。例如，Fortinsky 等发现行为障碍的严重程度与照料者症状管理自我效能和社区支持服务的利用自我效能水平呈负相关，患者的认知损害程度与社区支持服务的利用自我效能水平呈负相关，但未发现 ADLs 和 IADLs 与特定的自我效能间的相关性。Gottlieb 则报道未发现行为障碍与关系、工具和自我安抚这三个特定的自我效能水平的相关性。

在照料者自我效能的研究中，社会支持被认为是重要的保护性资源。尽管采用不同的自我效能和社会支持工具，研究发现照料者自我效能感水平与社会支持呈显著正相关。采用《修订版照料自我效能量表》进行研究发现三个维度的自我效能水平即获得休息、回应患者的困扰行为和控制照料相关懊恼想法，与社会支持呈显著正相关。笔者对上海地区痴呆患者家属照料者的调查发现，照料者积极的社交互动水平越高，照料者在收集照料信息、应对患者 BPSD 和压力管理方面的自我效能感也越强，同时照料者的精神健康水平也越高。

在照料负担研究中，社会支持被认为是一个调节因子。Bandura（1997）认为社会支持通过自我效能的中介作用影响人们的健康行为。Au 等报道自我效能对社会支持与照料者抑郁症状间存在部分中介作用。笔者对上海地区痴呆患者配偶照料者的研究发现，照料者不同行为维度的自我效能感对其社交受限程度与精神健康水平间的关系有部分中介作用。例如，收集照料信息和照料相关的压力管理自我效能感对积极社交互动与照料者精神健康间的关系有部分中介作用。笔者的研究还发现，配偶照料者的自我效能感是社会支持

和照料者主观负担间的调节因子，社会支持也是自我效能感和照料者主观负担间的调节因子。虽然照料者自我效能的研究越来越受到痴呆家属照料领域研究者的关注，但基于自我效能理论的相关路径分析的研究仍然很有限。

二、自我效能与照料者健康结果

尽管大量的文献报道了照料痴呆患者导致家属照料者不良健康的结果，但照料者自我效能与其健康结果的研究较有限。抑郁症状是最常见的与照料相关的不良健康结果，也是大多数照料者自我效能研究关注的结果指标。采用不同自我效能量表的研究发现照料者自我效能与抑郁症状呈负相关。研究发现以下 3 种自我效能——症状管理和自我照顾的自我效能、获得休息与回应患者的困扰行为、控制照料相关懊恼想法——水平高的照料者的抑郁症状明显较轻；症状管理自我效能对患者神经精神症状和照料者抑郁症状之间关系存在中介作用；面对患者行为障碍时，控制照料相关懊恼想法自我效能水平高的照料者照料负担较轻。研究还发现自我效能水平高的照料者对照料有较多积极的想法，消极情绪反应如焦虑和愤怒也较少。此外，笔者的研究发现不同行为维度的自我效能感对其社交受限程度与精神健康水平间的关系存在部分中介作用，如收集照料信息和照料相关的压力管理自我效能感对积极社交互动与照料者精神健康间的关系有部分中介作用。

主观负担也是最常见的结果评价指标。研究发现照料者自我照顾和照料自我效能以及解决问题自我效能水平与主观负担呈显著负相关。症状管理自我效能与主观负担呈显著负相关。

相对于精神心理相关的结果指标，自我效能与照料者躯体健康关系的研究非常有限，并且研究结果也不一致。Haley 等和 Wijngaart 等均未发现自我效能与照料者自评躯体健康间的关系。Wijngaart 等推测这一结果可能与两者反映的是照料者对自身能力评价的不同方面有关。照料者自我效能反映的是有效面对照料过程中困难的信心，而自评躯体健康反映的是其执行日常生活的能力。与之相反的是，Fortinsky 等报道照料者症状管理和社区支持服务利用自我效能感高的照料者躯体健康水平较高。Gottlieb 和 Rooney 采用 SF－36 评价照料者自我效能与躯体健康的关系，发现自我安抚自我效能与躯体功能水平呈正相关，与躯体疼痛呈负相关。此外，对照料者应对自我效能（coping self-efficacy）的研究发现，控制年龄、性别、烟酒史、BMI、糖尿病史、服用抗高血压药物史和痴呆严重程度等影响因素后，积极应对的自我效能较高

的照料者平均动脉压、收缩压和脉压较低；配偶照料者应对自我效能对照料负担和其 IL-6 值间有交互作用，在应对自我效能较低的照料者中照料负担与 IL-6 值显著相关，而在应对自我效能较高的照料者中未见这一显著相关性。

照料者的健康相关生活质量(health-related quality of life，HRQoL)一直是痴呆照料研究中的重点，负面的情绪、主观负担、躯体症状均为 HRQoL 的重要方面，但有关照料者自我效能与 HRQoL 的研究较有限。Gottlieb 采用《RIS 老年照护自我效能量表》和常用评价照料者 HRQoL 的工具 SF-36 进行的队列结果发现 RIS 的 3 个分量表分(工具、关系和自我安抚自我效能)均与 SF-36 的精神健康与活力分量表分呈显著正相关。然而，笔者对照料者不同照料行为维度的自我效能感与其健康相关结果(主观照料负担、精神健康和生活质量)间关系的研究发现，管理患者日常生活的自我效能感水平越高，照料者的精神健康水平越低，提示家属照料者对患者日常生活照料的信心越高。这一结果进一步提示对照料者自我效能及其影响因素的研究必须涵盖该群体主要照料行为维度的自我效能感，而非一或两个维度。笔者的研究还发现照料者自我效能感对其感受到的社会支持与其健康相关生活质量间关系有部分中介作用。

此外，已报道的一些将照料者自我效能感作为二级结果指标的干预研究并没有依照自我效能的理论模型选择干预形式，而是简单地将该指标列入各自的研究，从而导致结果不一致，因此还需要进一步的研究确定体现该结果指标的敏感性的干预形式。

三、自我效能评估

作为较新的评价变量，目前的研究所采用的量表包括《照料者自我效能量表》《照料者自我效能问卷》《RIS 老年照护自我效能量表》《激惹管理自我效能量表》《修订版照料自我效能量表》，以及《中国家属照料者自我效能问卷》(self-efficacy questionnaire for Chinese family caregivers，SEQCFC)(表 13-3-1)。这些自我效能量表的评价维度不同，但均包含了患者行为障碍管理和寻求相关帮助的维度。但其中大部分量表都非针对痴呆照料者群体而研发，有报道的国内外针对痴呆照料者照料行为的自我效能量表仅有两个，即《修订版照料者自我效能量表》和笔者研发的针对国内痴呆照料者的《中国家属照料者自我效能问卷》(SEQCFC)(Cronbach's α 均>0.8)。根据 Bandura 的自我

表 13-3-1　照料者自我效能量表

名称	条目、维度和评分	心理评估特性	备注
照料者自我效能量表	1. 共 14 个条目 2. 2 个维度：自我照顾(CSCE)和解决问题(CPSSE) 3. 每条目按 0～100 分评分，分量表总分为 0～100 分，得分越高，自我效能量水平越高	1. Cronbach α：CSCE，0.76；CPSSE，0.83 2. 11 周重测信度：CSCE，0.675；CPSSE，0.683 3. 会聚效度 社会支持：CSCE，$r=0.30$，$P<0.0001$；CPSSE，$r=0.19$，$P<0.05$ 应对：CPSSE，$r=0.19$，$P<0.05$； 抑郁症状：CSCE，$\beta=-0.262$，$P<0.001$；CPSSE，$\beta=-0.322$，$P<0.001$； 总体困扰：CSCE，$\beta=-0.111$，$P<0.05$；CPSSE，$\beta=-0.262$，$P<0.01$； 主观负担：CPSSE，$\beta=-0.202$，$P<0.01$； 4. 判别效度：应对，CSCE，$r=-0.04$。	1. 并非针对痴呆家属照料者而研制 2. 部分条目存在天花板效应
照料者自我效能问卷	1. 9 个条目 2. 2 个维度：症状管理(SXEFF)和社区支持服务的利用(SERVEFF) 3. 每条目按 1～10 分评分，得分越高，自我效能量水平越高	1. 结构效度：探索性因子分析产生两因子(SXEFF & SERVEFF) 2. Cronbach α：SXEFF，0.77；SERVEFF，0.78 3. 会聚效度 患者行为障碍：SXEFF，$r=-0.22$，$P\leqslant0.01$；SERVEFF，$r=-0.17$，$P\leqslant0.05$ 患者认知状态：SERVEFF，$r=-0.15$，$P\leqslant0.05$ 照料能力：SXEFF，$r=0.49$，$P\leqslant0.01$；SERVEFF，$r=0.27$，$P\leqslant0.01$； 照料者抑郁症状：SXEFF，$\beta=-0.17$，$P<0.05$； 照料者躯体症状：SXEFF，$\beta=-0.16$，$P<0.05$；SERVEFF，$\beta=-0.20$，$P<0.01$	1. 并非针对痴呆家属照料者而研制 2. 未检验重测信度

（续表）

名称	条目、维度和评分	心理评估特性	备注
RIS 老年照护自我效能量表	1. 10 个条目 2. 3 个维度：工具、关系和自我安抚 3. 每个条目按 0～4 分 5 级评分，得分越高，自我效能量水平越高	1. 结构效度：探索性因子分析产生三因子（工具、关系和自我安抚） 2. Cronbach α：工具，$\alpha_1=0.72$ 和 $\alpha_2=0.73$；关系，$\alpha_1=0.74$ 和 $\alpha_2=0.78$；自我安抚，$\alpha_1=0.79$ 和 $\alpha_2=0.72$ 3. 4～6 个月重测信度：工具，$\gamma=0.69$，$P<0.001$；关系，$\gamma=0.48$，$P<0.001$；自我安抚，$\gamma=0.60$，$P<0.001$ 4. 同时效度 SF-36 a）精神健康分量表：工具，$\gamma=0.19$，$P<0.05$；关系，$\gamma=0.30$，$P<0.01$；自我安抚，$\gamma=0.28$，$P<0.01$ b）活力分量表：工具，$\gamma=0.18$，$P<0.05$；关系，$\gamma=0.28$，$P<0.01$；自我安抚，$\gamma=0.29$，$P<0.01$ 人格特质 a）乐观：工具，$\gamma=0.28$，$P<0.01$；关系，$\gamma=0.41$，$P<0.001$；自我安抚，$\gamma=0.36$，$P<0.001$ b）随和：工具，$\gamma=0.31$，$P<0.01$；关系，$\gamma=0.22$，$P<0.05$；自我安抚，$\gamma=0.25$，$P<0.05$ c）尽责：工具，$\gamma=0.33$，$P<0.01$；关系，$\gamma=0.40$，$P<0.001$；自我安抚，$\gamma=0.29$，$P<0.01$ 愤怒情绪：工具，$\gamma=-0.29$，$P<0.01$；关系，$\gamma=-0.35$，$P<0.001$；自我安抚，$\gamma=-0.22$，$P<0.05$ 5. 会聚效度 社会支持：工具，$\gamma(134)=0.20$，$P<0.01$；关系，$\gamma(134)=0.23$，$P<0.001$；自我安抚，$\gamma(134)=0.30$，$P<0.001$ 应对效能：$\gamma(134)=0.26$，$P<0.01$ 至 $\gamma(134)=0.32$，$P<0.001$	并非针对痴呆家属照料者而研制
激惹管理自我效能量表	1. 43 个条目 2. 4 个维度：无身体攻击行为，身体攻击行为，无语言攻击行为，语言攻击行为	1. 表面效度 2. Cronbach α：总量表，0.86；各分量表（0.59～0.90）	1. 并非针对痴呆家属照料者而研制 2. 未做其他效度和重测信度检验

（续表）

名称	条目、维度和评分	心理评估特性	备注
修订版照料自我效能量表	1. 15个条目 2. 3个维度：获得休息（SE：OR），回应患者的困扰行为（SE：RDB）和控制照料相关懊恼想法（SE：CT） 3. 每条目按0～100分评分，分量表总分为0～100分，得分越高，自我效能量水平越高	1. 结构效度：探索性因子分析和验证性因子分析产生三因子（SE：OR；SE：RDB和SE：CT） 2. Cronbach α：SE：OR，0.85；SE：DB，0.82；SE：CT，0.85 3. 2周重测信度：SE：OR，0.76；SE：DB，0.70；SE：CT，0.76 4. 会聚效度和判别效度 抑郁症状：SE：OR，$\gamma=-0.38$，$P<0.001$；SE：DB，$\gamma=-0.34$，$P<0.001$；SE：CT，$\gamma=-0.38$，$P<0.001$ 焦虑：SE：OR，$\gamma=-0.24$，$P<0.01$；SE：DB，$\gamma=-0.20$，$P<0.05$；SE：CT，$\gamma=-0.37$，$P<0.001$ 愤怒：SE：OR，$\gamma=-0.04$，N；SE：DB，$\gamma=-0.45$，$P<0.001$；SE：CT，$\gamma=-0.31$，$P<0.001$ 社会支持：SE：OR，$\gamma=0.47$，$P<0.001$；SE：DB，$\gamma=0.08$，N；SE：CT，$\gamma=0.10$，N 5. 同时效度：低批判组的照料者（$n=58$）SE：DB和SE：CT分高于边界和高批判组（$n=34$）	1. 在Zeiss的量表基础上，针对痴呆家属照料者而修订 2. 条目内容反映了痴呆照料者的特定照料行为
中国家属照料者自我效能问卷	1. 27个条目 2. 5个维度：收集照料信息（GI），获取帮助（OS），回应患者行为障碍（RBD），对患者日常生活的管理（MHPMC），照料相关的压力管理（MDC） 3. 每条目按0～100分评分，量表及分量表总分为0～100分，得分越高，自我效能量水平越高	1. 结构效度：探索性因子分析产生5因子 2. Cronbach α：量表几个分量表均>0.8 3. 四周重测信度：0.64～0.85 4. 会聚效度 社会支持：GI，$\beta=0.25$，$P<0.001$；OS=0.42，$P<0.001$；RBD，$\beta=0.28$，$P<0.001$；MDC，$\beta=0.23$，$P<0.001$ 认知功能：RBD，$\beta=-0.24$，$P<0.05$ 躯体功能：MHPMC，$\beta=0.33$，$P<0.01$ 行为障碍：RBD，$\beta=-0.22$，$P<0.05$；MHPMC，$\beta=-0.31$，$P<0.001$；MDC，$\beta=-0.30$，$P<0.001$	1. 中国痴呆家属照料者而研制 2. 条目内容反映了痴呆照料者的特定照料行为

效能理论，自我效能的广度只有与特定目标行为密切相连，才能有效预测个体的行为；且自我效能水平的变化及个体的行为还受个体所处的社会、经济、文化等情境因素的影响。简而言之，评价照料者自我效能感的工具必须能体现照料者在所处的社会文化背景下的常见照料行为。国内外文献表明，痴呆患者家属照料者最常见的照料行为包括寻求帮助或获得休息、管理 BPSD、照料压力的管理等。笔者对上海地区痴呆患者家属照料者照料行为的质性研究发现中国家属照料者有五大类常见照料行为，即收集照料信息、获取帮助、回应患者 BPSD、对患者日常生活的管理，以及照料相关的压力管理。SEQCFC 是在该质性研究的基础上，借鉴了《修订版照料者自我效能量表》的部分条目发展而成的。特别指出的是，《修订版照料者自我效能量表》和 SEQCFC 均以 Bandura 自我效能理论为框架，条目内容反映了痴呆照料者的特定照料行为，笔者在研制 SEQCFC 的同时也完成了中国痴呆患者照料者自我效能理论模型的初步测定。虽然上述量表的信效度是可接受的，但都需要更广泛的研究，特别是干预性研究进一步验证。

（张曙映　郭起浩）

第四节　阿尔茨海默病相关结果评价量表（ROSA）

阿尔茨海默病相关结果评价量表（relevant outcome scale for Alzheimer's disease, ROSA）评定是对于阿尔茨海默病（AD）病程严重程度基于临床表现进行的反复评估。它作为测量工具被开发用于临床实践中，可以由受过训练的医师和其他人员进行评估（如经过培训的心理测验评定员）。

ROSA 估计疾病程度时，按照疾病的严重程度（早、中、晚 3 个阶段），使用了 14 个项目，并设计了患者生活质量的评估和照料者生活负担评估，共 16 个项目。患者能力评估涉及 4 个方面：① 认知（项目 1～3）；② 沟通（项目 4～6）；③ 行为（项目 7～11）；④ 日常生活活动（项目 12～14）。这 14 个项目都是日常情景描述。评估方法是相同的，按照严重性进行分段来区分疾病的严重程度和预测患者的能力/行为；评估是依据患者的能力/行为在数值范围 10 到 0 之间选相应的数值（如很好、很差）。最终，评估者得出总分并给予患者相应的评价结果。

ROSA 以访谈的形式进行。在疾病的早期阶段，可与患者访谈。同时，一

般情况下，患者的陪同人员因经常和患者在一起，他们可以回答一些常见的问题。该评估的最终结果由测量者根据评估结果、患者的既往史、临床数据并结合测量者个人经验获得。使用 ROSA 的一个条件是，用户必须熟悉必要的背景资料，确认患者可以作为评估对象。

如何使用 ROSA？首先，评定员须对 AD 患者进行全面临床评估。这一评估的基础是必须对患者的临床表现准确判断。另外，要会使用评估量表，如 GDS 可用来确定患病阶段。但是，这不是一个强制性的要求。

阿尔茨海默病可分早、中、晚三期。早期，患者表现不明显，亲人、朋友或同事一般不易觉察到患者的认知障碍。认知障碍包括：词和名字难以被想起，记忆反复出现最近经历的事情或过往生活中的琐事。患者可能会在计划及执行社会角色方面出现困难（如家庭聚会、结识新朋友）。中期，近事健忘和已往记忆增加，完成复杂事物有困难（如策划活动、写信件、支付账单等），抑郁，退出社交，或有其他行为的明显改变；其日常生活可能发生障碍，如根据季节选择合适的服装或准备食物不能自理。晚期，会发生患者失忆事件，忘记自己是谁，日常生活不能自理（如饮食和如厕），空间和时间定向严重受损，依据季节选择合适的衣服需要帮助，个性和行为变化（如冷漠、抑郁、妄想、不安），以及夜间睡眠模式发生改变，反应能力严重受损。

在 ROSA 中，提出了 14 个情景。对于疾病每个阶段严重性依据患者的能力表现都有 3 个程度的描述。评估患者的能力或行为应符合患者的实际能力，运用量表中从 10 到 0 的数值进行标记。在评估时，每个问题只能有一个值被标记。使用 ROSA 前确定每个阶段患者的严重程度。请注意，患者疾病的严重程度是动态的，这仅仅是在当下评估的患者的能力/行为。

请记住，每一种情况下患者的能力和行为在最后一周应评估。评价应结合情景帮助以及参考相关例子。如果没有为患者得出一个恰当的结论，则应请教相关领域的专家。这些都必须加以记录（如一个附加页），进一步重复测试，并且保持不变。例如：

场景 3：患者会喜欢去购物/准备食物。其他：患者要挂一幅画/修剪草坪/种花/启动工艺品项目。

场景 8：患者安静地坐着。其他：向大家介绍孙子就读的学校/音乐/剧院/与朋友聚餐的场景。

使用 ROSA 的步骤如下：

(1) 基于不同阶段的阿尔茨海默病（早期、中期、晚期）全球评估程度的一

致性。请用×表明严重程度。

(2) ROSA 量表 1～14。阅读相关情景/例子，这个量表可反映阿尔茨海默病（早、中、晚期）的程度。继续评估患者：标记数值尺度。这样的场景评估可以推断阿尔茨海默病的严重程度。该说明中提到个别不适用的患者，可使用合适的替代方案。目前提供的这张纸，可作为个别患者进一步评估的模板，以便更好地处理患者的情况。差或不足数值代表最低点为 0，最好的表现是 10。

(3) 最后两项是对患者生活质量和照顾者负担的评估，需要结合评估者在前 14 项所做的 ROSA 评估的基础上对患者的整体状况进行评价。

(4) ROSA 总分是输入所有标明的数值及有关项目的得分框旁边的得分而获得的。总得分可反映每个患者的严重程度，从 0～160。除了总分，为了准确评估 AD 患者的病情，也可以编写相应的项目分数值标注在量表之后。

一、重复评估

如果是同一种疾病同一病程做的 ROSA 评估结果可以直接比较。如果患者病情出现变化，ROSA 评估结果可以和另一个患者对照。请记住，一般 ROSA 评估必须始终在相同条件下进行，最好是相同的评估者，这样才能保证评估结果的相对一致性。请注意每一个人的偏差（如情景、采访合作伙伴）。亲戚或照料者在参与评估过程中，应尽可能使家属自由地表达自己的感受。

二、ROSA 的优缺点

1. 优点　① 全面性：不仅覆盖了 ADAS - cog 反映的记忆、语言、执行这些基本认知功能领域，DAD 反映的日常生活能力，NPI 反映的神经精神症状，还包括社交活动、配合性、兴趣、生活质量等更高层次的评估。② 得分范围大，便于随访观察严重度变化。③ 简洁、易携带、耗时少、易推广。

2. 缺点　① 主观性：在 10 个等级中，要准确区分级别是有难度的，尤其是没有经过严格训练的评定员。② 总体严重度与项目严重度的不一致：量表要求首先判断总体严重度，分为轻、中、重度，再进一步区分等级。有些项目，如攻击性、坐立不安、配合性，未必是晚期就比中期严重。③ 不典型 AD，如额叶型 AD、后部皮质萎缩（PCA）、少词性进行性失语（logopenic PPA），ROSA 的适用性尚不清楚。

我们已经完成中文版 ROSA（CROSA）（表 13 - 4 - 1）的信度、效度的验证。结果表明，CROSA 总分和其各项目分的相关性较好（相关系数基本为

表 13-4-1　阿尔茨海默病相关结果评价量表(ROSA)

患者姓名：________　评分员姓名：________　日期：日 月 年

我们希望您能告诉我们，您所照顾的患者所患疾病对其应对日常生活的能力的影响。请在最佳描述患者前一周能力的评价上打叉（×）。请尽量回答所有的问题，即使该问题目前不适用。没有对或错的答案。您的回答应准确反映您对该患者能力的评价。谢谢！

该患者的阿尔茨海默病处于　□早期　□中期　□晚期

评分 | 权重 | 得分

1. 该患者能够回忆起很久以前的事情。

非常好 □10 □ □9 □ □8 □ □7 □ □6 □ □5 □ □4 □ □3 □ □2 □ □1 □ □0 非常差/根本不　← ×1 =

场景：该患者遇到一位至亲或好朋友。
早期：该患者认识这位亲戚/朋友，但不知道他的确切姓名。
中期：该患者难以记起这位亲戚/朋友的姓名以及与他的关系。
晚期：该患者难以记起这位亲戚/朋友的姓名以及与他的关系。该病人难以记起他自己的姓名。

2. 该患者能够回忆起最近的事情。

非常好 □10 □ □9 □ □8 □ □7 □ □6 □ □5 □ □4 □ □3 □ □2 □ □1 □ □0 非常差/根本不　← ×1 =

场景：该患者想出去走走，但想不起钥匙放在哪里。
早期：该患者找了几分钟钥匙。该患者没有找到钥匙，也就没有出门。
中期：该患者知道应该怎么办，但还是没有带钥匙就出门了。
晚期：大多情况下，该患者不记得他出门时应该带着钥匙。

3. 该患者能够计划并执行复杂的事情。

非常好 □10 □ □9 □ □8 □ □7 □ □6 □ □5 □ □4 □ □3 □ □2 □ □1 □ □0 非常差/根本不　← ×1 =

场景：该患者想出去购物/准备食物。
早期：该患者能够在少许帮助或指导下进行购物/准备食物。
中期：该患者购物有很大的难度，但在帮助下仍然能够准备简单的饭菜。
晚期：即便是简单的事情该患者也需要帮助。

4. 该患者使他人理解的能力。

非常好 □10 □ □9 □ □8 □ □7 □ □6 □ □5 □ □4 □ □3 □ □2 □ □1 □ □0 非常差/根本不　← ×1 =

场景：该患者想告诉别人一些事情。
早期：该患者需要一些时间找到确切的词语。
中期：该患者不能使用确切的词语表达他的意思。他人难以理解该患者想说什么。
晚期：即使该患者仍然能够说话，但难以使他人理解。

5. 该患者的沟通能力。

非常好 □10 □ □9 □ □8 □ □7 □ □6 □ □5 □ □4 □ □3 □ □2 □ □1 □ □0 非常差/根本不　← ×1 =

场景：该患者想让别人为他做些事情。
早期：该患者常常清楚、确切地表达他的意愿，并能理解他人的回应。
中期：该患者难以选择确切的词语有条理性地表达他的意愿，难以理解他人的回应。
晚期：该患者通过声音或手势表达他的意愿，几乎不能理解他人的回应。

6. 该患者表现出社交能力。

非常好 □10 □ □9 □ □8 □ □7 □ □6 □ □5 □ □4 □ □3 □ □2 □ □1 □ □0 非常差/根本不　← ×1 =

场景：该患者想接触他人。
早期：该患者有时找人与他度过一段时间，或打电话给朋友/家人，以使他们能够与他度过一段时间。
中期：该患者要在别人的鼓励下接触他人/与他人交谈。
晚期：大多情况下，该患者没有反应，即使有人想与他接触。

7. 该患者行为的攻击性。

根本不 □10 □ □9 □ □8 □ □7 □ □6 □ □5 □ □4 □ □3 □ □2 □ □1 □ □0 非常多　← ×4 =

场景：该患者想要一些东西，但没有得到/该患者想要别人为他做些事情，但别人没有做。
早期：该患者用言辞表达对此情形的愤怒。
中期：在这样的情形下，该患者大喊/咒骂或带有威胁性的手势。
晚期：在这样的情形下，该患者咒骂或常常变得暴力。

＋

中间合计（I）：

	评分	权重	得分
8. 该患者坐立不安。 根本不 □ 非常多 10 9 8 7 6 5 4 3 2 1 0 场景：该患者安静地坐着。 早期：没有任何明显的触发，该患者开始自言自语或轻微摇晃身体。 中期：该患者突然开始转来转去。 晚期：该患者不能安静地坐着，他长时间的转来转去并自发性地咒骂。	←	×4=	
9. 该患者因妄想而表现出行为上的变化。 根本不 □ 非常多 10 9 8 7 6 5 4 3 2 1 0 场景：该患者表现出妄想的征象/存在妄想，认为有人偷他的钱或偷他的家或老伴不属于他。 早期：该患者确信有人偷他的钱。 中期：该患者对“错误的”老伴表现出愤怒，或试图离开这个“错误的”家。 晚期：该患者对“错误的”老伴反应激烈，或试图阻止他离开这个“错误”的家。	←	×4=	
10. 该患者的不安全感。 根本不 □ 非常多 10 9 8 7 6 5 4 3 2 1 0 场景：该患者不想独自待在不寻常的场合。 早期：该患者独处时，有时表现出不安或焦虑。 中期：该患者独处时，常常表现出紧张、不安或焦虑。 晚期：该患者常常非常焦虑，开始哭或试图抓住他人不放。	←	×4=	
11. 该患者行为的配合性。 非常多 □ 根本不 10 9 8 7 6 5 4 3 2 1 0 场景：假定该患者去看医生/睡觉。 早期：该患者不喜欢非得去看医生/睡觉，但还是做了。 中期：该患者一定要在劝说下才去看医生/睡觉。 晚期：该患者强烈拒绝去看医生/睡觉，也几乎不能被劝说去这样做。	←	×4=	
12. 该患者能够胜任每天的日常生活。 非常多 □ 不多/根本不 10 9 8 7 6 5 4 3 2 1 0 场景：该患者想穿衣服/吃东西。 早期：该患者能够挑选衣服/食物，能够自己穿衣服/吃东西。 中期：该患者能够穿上为他放好的衣服/自己吃准备好的饭菜。 晚期：该患者自己穿衣服/吃东西有很大的困难。	←	×1=	
13. 该患者能够留心周边环境，并表现出兴趣。 非常多 □ 非常少 10 9 8 7 6 5 4 3 2 1 0 场景：该患者对日常事件感兴趣（家庭/体育运动/股票交易等）。 早期：该患者在电视/广播上收看收听日常事件（等），并与他人谈论。 中期：该患者在被问及时会谈论日常事件（等）。 晚期：即使该患者被告知一些日常事件（等），他常常没有反应。	←	×1=	
14. 该患者能够自理。 非常多 □ 根本不 10 9 8 7 6 5 4 3 2 1 0 场景：该患者想去一些地方或做一些事情。 早期：该患者自己决定想去哪里、想做什么/仅需少许帮助。 中期：该患者需要在鼓励和帮助下去一些地方/做一些事情。 晚期：即使有他人陪伴，该患者也几乎不能去一些地方/做一些事情。	←	×1=	
15. 该患者的生活品质。 非常好 □ 非常差 10 9 8 7 6 5 4 3 2 1 0	←	×1=	
16. 照顾者的负担 非常小 □ 非常大 10 9 8 7 6 5 4 3 2 1 0	←	×1=	
			+
中间合计（2）：			
中间合计（1）：			
总分范围：0~310　　总分：			

0.4～0.7)；CROSA 总分和常用的评估阿尔兹海默病认知及生活能力受损的量表总分的相关度较高(MMSE, DRS, ADAS－cog, IQCODE, DAD, NPI, ZBI)，相关系数基本为 0.4～0.7(除与 ZBI 的相关系数为 0.3)；且 CROSA 对语言、认知等各领域评估的分数与侧重评估该种认知功能的量表的相关度也较好(如 CROSA 的认知、交流和 ADAS－cog、IQCODE 的相关度较高，行为和 ZBI 的相关度较高，等等)。说明 CROSA 评估有较好的信度和效度。

此外，CROSA 也可以快速地、较好地反应药物治疗的效果。在经过 12 周的治疗后，CROSA 总分和某些项目分，都显著地升高，且在行为、生活质量、照料者负担方面都可以看到显著的改善，这是传统的评估中涉及相对较少的方面。

(杨　青　郭起浩)

第五节　日常认知量表(ECog)

日常认知量表(everyday cognition, ECog)是由 Tomaszewski Farias S 等于 2008 年编制的，是知情者日常认知评估量表，用于检测认知功能减退，包含 39 个项目，覆盖 6 个认知相关域：日常记忆、日常语言、日常视觉空间能力、日常计划、日常组织和日常注意分割(everyday divided attention)(表 13－5－1)。对于每一项，知情者都将参与者当前的日常功能水平与他(或她)10 年前的比较。个体作为他们自己的对照。

表 13－5－1　ECog 长版(39 项 ECog 版本)

1＝比 10 年前的状况要好，或跟 10 年前差不多；2＝好像差一点，或者偶尔差一点；3＝肯定差一点；4＝差多了，或明显减退；5＝不知道，无法判断，或者患者从不做该事。

项　目	严重度	
1. 无须列出清单就能记住所有需要购买的东西	1 2 3 4	5
2. 记得最近发生的事情，如近期的外出、最近的新闻事件等	1 2 3 4	5
3. 记得几天以前的谈话内容	1 2 3 4	5
4. 记得自己把东西放在哪里	1 2 3 4	5
5. 能复述故事，或能记得自己提过的问题	1 2 3 4	5
6. 记得今天是几号或星期几	1 2 3 4	5
7. 记得自己曾经和某人说过某件事	1 2 3 4	5

（续表）

项　　目	严	重	度		
8. 能记得预约、会议或债务	1	2	3	4	5
9. 能给其他人下口头指令	1	2	3	4	5
10. 能读懂书中的一个故事	1	2	3	4	5
11. 能听懂其他人谈话的要点	1	2	3	4	5
12. 能描述他/她在电视里看过的一个节目	1	2	3	4	5
13. 能听懂口头的指令	1	2	3	4	5
14. 忘记东西的名称	1	2	3	4	5
15. 在谈话中用词得体	1	2	3	4	5
16. 能记得常用词语的意思	1	2	3	4	5
17. 在与人交谈中能正确表达自己的想法	1	2	3	4	5
18. 能看着地图找到一个新地方	1	2	3	4	5
19. 在其他人开车时，能看地图帮忙指路	1	2	3	4	5
20. 能在停车场找到自己的车	1	2	3	4	5
21. 在商场或其他地方，能找到路去约定好的地点与他人会合	1	2	3	4	5
22. 在熟悉的地方能认得路	1	2	3	4	5
23. 在熟悉的商场里能认得路	1	2	3	4	5
24. 到去过很多次的地方时，能认得路	1	2	3	4	5
25. 在买东西时能合理安排停留地点	1	2	3	4	5
26. 能预计天气的变化，并根据变化调整自己的计划	1	2	3	4	5
27. 对可能要发生的事情预先做好应对计划	1	2	3	4	5
28. 能展望未来，为今后的事情做打算	1	2	3	4	5
29. 在做事之前先把过程思考一遍	1	2	3	4	5
30. 保持工作和生活环境的整洁	1	2	3	4	5
31. 准确记账而不出错	1	2	3	4	5
32. 保持财务记录的条理性	1	2	3	4	5
33. 能根据重要性的不同安排事情的先后顺序	1	2	3	4	5
34. 到医院看病时，能安排好次序以节约时间，避免走弯路	1	2	3	4	5
35. 能整理信件和报纸	1	2	3	4	5
36. 能同时做两件事	1	2	3	4	5
37. 在做事的中途被打断后，还能继续做原来进行的事情	1	2	3	4	5
38. 能专心做一件事，不被外界环境所干扰	1	2	3	4	5
39. 烧饭或工作的同时能说话	1	2	3	4	5

2011 年 Tomaszewski Farias S 开发并验证了 ECog 简短版(short form)即 12 项 ECog 版本(ECog-12)(表 13-5-2),其在 ECog 6 个认知相关域中的每个域各挑选出两项,每个域中被挑选出的那两项分别具有最高和最低的项目难度评估。ECog-12 被证实与完整的 39 项版本有极强的相关性,与年龄和教育弱相关,具有高度的内部一致性。ECog-12 能够很好地识别痴呆人群与正常人群[ROC 曲线下面积(AUC)=0.95,CI=0.94~0.97],并有希望将正常年长者从认知受损者(MCI 或痴呆)中识别出来。但是识别 MCI 与正常人群的能力较弱。

表 13-5-2 ECog 简短版(12 项 ECog 版本)

1=比 10 年前的状况要好,或跟 10 年前差不多;2=好像差一点,或者偶尔差一点;3=肯定差一点;4=差多了,或明显减退;5=不知道,无法判断,或者患者从不做该事

项目	严重度				
1. 记得自己把东西放在哪里	1	2	3	4	5
2. 记得今天是几号或星期几	1	2	3	4	5
3. 能听懂口头的指令	1	2	3	4	5
4. 在与人交谈中能正确表达自己的想法	1	2	3	4	5
5. 能看着地图找到一个新地方	1	2	3	4	5
6. 到去过很多次的地方时,能认得路	1	2	3	4	5
7. 能预计天气的变化,并根据变化调整自己的计划	1	2	3	4	5
8. 对可能要发生的事情预先做好应对计划(未雨绸缪)	1	2	3	4	5
9. 保持工作和生活环境的整洁	1	2	3	4	5
10. 准确、有条不紊地记账而不出错	1	2	3	4	5
11. 能同时做两件事(如一边看电视一边打毛线)	1	2	3	4	5
12. 烧饭或工作的同时能说话	1	2	3	4	5

ECog 评分方法将患者认知功能改变程度分为 4 个等级:1=比 10 年前的状况好,或跟 10 年前差不多;2=好像差一点,或者偶尔差一点;3=肯定差一点;4=差多了,或明显减退。

将所有完成项目的得分相加得到一个总分,然后除以完成的项目数,总体得分范围为 1~4。ECog-39 得分通过计算所有 39 个项目的平均分获得,同样地 ECog-12 通过计算 12 个项目的平均分获得。

信度:内部一致性通过计算 Cronbach's α 系数来测量。ECog-12 的 Cronbach's α 系数为 0.96,支持其作为一种通用的日常认知评估。ECog 量表具有良好的重测信度($r=0.82$)。

效度：ECog-39 与 ECog-12 在识别认知受损者(MCI 或痴呆)与认知功能正常者时同样敏感。ECog 两个版本的 ROC 曲线下面积为 0.91(95%CI=0.88～0.93)。当设定敏感性为 80%时确定最佳的特异性和划界分，ECog-39 与 ECog-12 的特异性分别为 84%和 81%，最佳划界分分别为 1.81 和 2.30。在识别痴呆与正常人群时，ECog-39 与 ECog-12 的特异性为 95%，最佳划界分分别为 2.34 和 2.83。在识别 MCI 与正常人群时，ECog-39 与 ECog-12 的特异性分别为 61%和 62%，最佳划界分分别为 1.32 和 1.54。ECog 量表还具有较好的会聚效度与分歧效度。

ECog 量表与人口统计学变量(年龄或教育水平)的相关性可以被忽略，ECog-39 和 ECog-12 与年龄的相关系数分别为 0.04 和 0.03，与教育的相关系数为 0.01，说明其不受患者年龄、教育水平的影响。

与 ECog 长版相比，ECog-12 在痴呆组中显示了"地板效应"，其痴呆组中 26%的得分在最高受损水平，而在 ECog 长版中仅为 5%。因此，临床医生或研究者是选择使用 ECog 简短版还是最初的长版依赖于评估目的及研究样本的特性。如果纵向追踪大范围的能力水平和疾病的严重度，最初的长版较合适。

知情者特征，如情绪或照料者负担程度会影响 ECog 评分分级。

Marshall 等发现 ECog 量表中与工具性日常生活能力(IADL)相关的一些项目可以最好地辨别临床正常个体(CN)与 MCI，并且预测从 CN 到 MCI 的进展。"无须列出清单就能记住所有需要购买的东西""能记得已经约好的会议、事情等""能展望未来，为今后的事情做打算""准确、有条不紊地记账而不出错""能整理信件和报纸"这五项较差表现最有助于将 MCI 从 CN 中辨别出来。"能整理信件和报纸"这一项较差表现可以有助于预测从 CN 到 MCI 的进展。

第六节　功能活动问卷(FAQ)

功能活动问卷(functional activities questionnaire, FAQ)由 Pfeffer 于 1982 年编制，是一个关于功能性 ADL 的知情者问卷，用于研究社区老年人的正常老化和轻度老年性痴呆，后于 1984 年进行了修订。FAQ 优于工具性日常生活能力(IADL)量表，运用 FAQ 区分正常与痴呆个体更敏感。

FAQ 原版有 11 个项目，采用 4 级评分，满分 33 分(表 13-6-1)。中文版最早由张明园教授修订，有 10 个项目，满分 30 分，删除了"收集纳税记录"(表

13-6-2)。FAQ主要评定完成项目的正确性：0=没有任何困难，能独立完成，不需他人帮助；1=有些困难，需要他人指导或监督；2=需要帮助才能完成，部分依赖；3=本人无法完成，或完全由他人代替完成。此问卷可由知情看护者在5 min内即可完成。

表13-6-1　FAQ英文直接翻译版

项　　目	评　分
1. 支付账单/平衡收支	0　1　2　3
2. 汇编税务记录	0　1　2　3
3. 自行购物(如购买衣服或杂货)	0　1　2　3
4. 参加技巧性的游戏(如桥牌或象棋)/从事业余爱好	0　1　2　3
5. 加热水和熄灭炉子	0　1　2　3
6. 准备一顿调配合理的饭菜	0　1　2　3
7. 追踪当前的事件	0　1　2　3
8. 集中注意力并能理解电视节目内容	0　1　2　3
9. 记得约定	0　1　2　3
10. 驾车或外出离开邻近地区	0　1　2　3

注：FAQ项目得分范围：0～3(得分越高提示功能受损越大；0=正常，或以前从不做但现在可以完成；1=有困难但可以自己完成，或以前从不做但现在完成有困难；2=需要帮助；3=依赖)。

表13-6-2　FAQ张明园修订版

说明：下面列举了10项活动，每项活动的评定分成以下几个等级： 0　没有任何困难，能独立完成，不需要他人指导或帮助 1　有些困难，需要他人指导或帮助 2　老人本人无法完成，完全或几乎完全由他人代替完成 9　不适用，如老人一向不从事这项活动 请仔细阅读(读出问题)，并按照老人的情况，做出能最合适地反映老人活动能力的评定，每一道问题只能选择一个评定，不要重复评定，也不要遗漏	
1. 使用各种票证(正确地使用，不过期)	0　1　2　9
2. 按时支付各种票据(如房租、水电等)	0　1　2　9
3. 自行购物(如购买衣、食及家庭用品)	0　1　2　9
4. 参加技巧性的游戏或活动(如打扑克、下棋、打麻将、绘画、摄影、集邮、练书法、做木工)	0　1　2　9
5. 使用炉子(包括生炉子和熄灭炉子)	0　1　2　9
6. 准备和烧一顿饭菜(有饭、菜或汤)	0　1　2　9
7. 关心和了解新鲜事物(如国家大事或邻居中发生的重要事情)	0　1　2　9
8. 持续一小时以上注意集中地看电视或看小说或听收音机，并能理解、评论或讨论其内容	0　1　2　9
9. 记得重要的约定(如领退休金、朋友约会、家庭事务和接送幼儿等)	0　1　2　9
10. 独自外出活动或走亲访友(指较远距离，如相当于公共车辆3站路的距离)	0　1　2　9

适用项目：　　　　总分：

资料可靠性：　1. 可靠　2. 基本可靠　3. 不大可靠

值得注意的是，时代变迁，各种功能活动也在变化，比如，通信工具由固定电话发展到移动电话，再到现在的智能手机，所以各个 FAQ 版本的条目有差异。

FAQ 只有正确性维度，笔者根据老年人功能缺损的特点，在已有的 FAQ "正确性"维度基础上增加了"完成速度"维度，其评分标准为：0＝与平时一样快；1＝完成该项目的时间比平时延长约 50%；2＝完成该项目的时间延长约 100%；3＝无限延长(表 13-6-3)。FAQ 根据完成项目的正确性与速度区分为 FAQ-Right 总分、FAQ-Speed 总分。由于 FAQ 原版本个别项目(如支付账单/平衡收支)不存在完成速度的问题，所以项目有调整。

表 13-6-3　FAQ 双维度评估

评分方法			
正　确　性		完　成　速　度	
0＝没有任何困难，能独立完成，不需他人帮助		0＝与平时一样快	
1＝有些困难，需要他人指导或监督		1＝完成该项目的时间比平时延长约 50%	
2＝需要帮助才能完成，部分依赖		2＝完成该项目的时间延长约 100%	
3＝本人无法完成，或完全由他人代替完成		3＝无限延长	
9＝该项目不适合，或老人从来不做		9＝无法完成该项目	
项　　目	**正确性**	**完成速度**	**不适合**
1. 使用电话或手机	0 1 2 3	0 1 2 3	9
2. 整理家庭物品井井有条、不凌乱	0 1 2 3	0 1 2 3	9
3. 自行购物(如购买衣服、食品及家庭用品)	0 1 2 3	0 1 2 3	9
4. 参加技巧性的游戏或活动(如打扑克、下棋、打麻将、绘画、摄影、集邮、练书法、做木工)	0 1 2 3	0 1 2 3	9
5. 使用各种电器(如电视、空调、微波炉、电饭煲)	0 1 2 3	0 1 2 3	9
6. 准备和烧一顿饭菜(包括加工蔬菜、使用炉子、调味品用量恰当)	0 1 2 3	0 1 2 3	9
7. 关心和了解新鲜事物(国家大事或邻居中发生的重要事情)	0 1 2 3	0 1 2 3	9
8. 持续一小时以上注意力集中地看电视或小说，或收听收音机并能理解、评论或讨论其内容	0 1 2 3	0 1 2 3	9
9. 记得重要的时间点(如领退休金日期、按时服药、接送幼儿等)	0 1 2 3	0 1 2 3	9
10. 独自外出活动或走亲访友(指较远距离、如相当于公共车辆 3 站路的距离)	0 1 2 3	0 1 2 3	9

FAQ 识别痴呆的敏感性和特异性比较高，与 MMSE 等客观评估联合应用可提高检测痴呆的性能。用于 MCI 识别的研究比较少，最近有一篇 Marshall 等的研究，发现 FAQ 中的一些项目可以很好地辨别临床正常个体

(clinically normal,CN)与MCI,并且预测从CN到MCI的进展。“记得重要的时间点”和“收集纳税记录”这两项最有助于区别CN与MCI。而“持续一小时以上注意力集中地看电视”“支付账单/平衡收支”“加热水和熄灭炉子”,这三项有助于预测从CN到MCI的进展。

(汤　乐　郭起浩)

第十四章

其他神经科量表

第一节　急性周围性面瘫的严重度评估量表

急性周围性面瘫在世界范围内都是一种很常见的神经系统疾病，美国报告年发病率约为 3 人/1 万并在逐年上升。面神经的解剖发布见图 14-1-1。Bell 面瘫呈急性起病、单侧、周围性面瘫。Hunt 综合征由膝状神经节病毒感染引起，表现面肌瘫痪外，还有舌前 2/3 味觉丧失，舌下腺、下颌下腺和泪腺等分泌障碍。为了有效评估面瘫的治疗效果、分析不同治疗方法（如针灸治疗、手术治

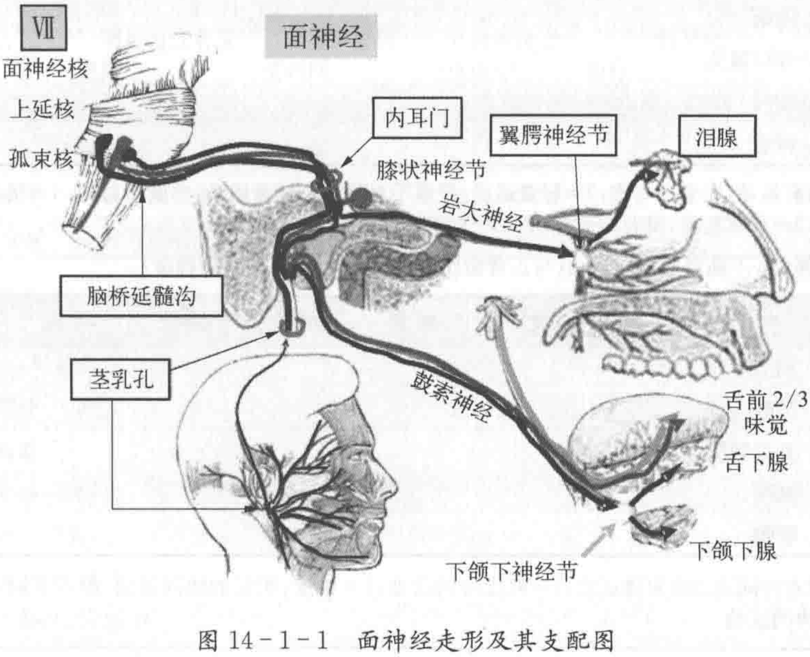


图 14-1-1　面神经走形及其支配图

疗和不同药物治疗)的疗效、寻找面瘫预后的判断指标,笔者根据面神经的解剖分布特点编制了面瘫评定系统(表 14-1-1、表 14-1-2)。同时,介绍国外常用的评估方法,这些方法临床使用存在评估不够全面、项目缺乏量化定义等不足。

表 14-1-1　急性周围性面瘫的严重度评估表

姓名　　年龄　　性别　　职业　　发病日期　左/右侧面部
联系电话　　　　地址
病前有无以下情况:(1) 受凉　(2) 糖尿病　(3) 高血压　(4) 颜面疱疹　(5) 长期吸烟　(6) 酗酒　(7) 颅底疾病　(8) 疲劳　(9) 心理压力过大　(10) 面瘫家族史　(11) 以往面瘫史　(12) 关节炎　(13) 其他

评定日期就诊时[　月　日]　10 天[　月　日]　20 天[　月　日]
30 天[　月　日]　40 天[　月　日]　50 天[　月　日]
治疗方法记录

Sunnybrook 面瘫评定系统

第一部分:安静状态下的面孔对称性(将病侧与正常侧比较)

部位	评分标准	就诊	10 天	20 天	30 天	40 天	50 天
眼睛	正常=0;缩小=1;扩大=1;眼睑手术=1						
鼻唇沟	正常=0;消失=2;变深=1;变浅=1						
嘴巴	正常=0;嘴角低垂=1;嘴角拉长=1						

第二部分:随意肌运动的对称性(与正常侧比较,患侧肌肉移动程度)

表　情	就诊	10 天	20 天	30 天	40 天	50 天
1. 前额皱纹						
2. 闭眼						
3. 开口微笑						
4. 喊叫						
5. 噘嘴						

1=没有运动,完全不对称;2=轻微运动,严重不对称;3=轻度运动,中度不对称;4=接近正常,轻度不对称;5=正常运动,没有不对称

第三部分:不随意肌协同运动(与正常侧比较,患侧不随意肌肉收缩程度)

表　情	就诊	10 天	20 天	30 天	40 天	50 天
1. 前额皱纹						
2. 闭眼						
3. 开口微笑						
4. 喊叫						
5. 噘嘴						

0=没有协同运动或集体运动;1=轻度协同运动;2=中度,明显的协同运动,但不丑陋;3=重度,变丑陋的协同运动

（续表）

Yanagihara 面瘫评定系统

表　情	就诊	10 天	20 天	30 天	40 天	50 天
1. 静止时不对称						
2. 前额皱纹						
3. 眨眼						
4. 轻闭眼						
5. 紧闭眼						
6. 闭受累单眼						
7. 鼻翼运动						
8. 开口微笑						
9. 吹口哨						
10. 紧闭嘴唇						

0＝不能（缺乏）；2＝下降，部分瘫痪；4＝正常

House Brackmann 面瘫评定系统

［就诊时　10 天　20 天　30 天　40 天　50 天］

级别	描述	总体	休息时	运动时		
				前额	眼睛	嘴巴
Ⅰ	正常	所有面部功能正常				
Ⅱ	轻度	细查细微无力	对称，紧张度相同	中度至接近正常	用力极小即能完全闭合	轻微不对称
Ⅲ	中度	明显但没有变丑陋	对称，紧张度相同	只能轻微至中度运动	用力能完全闭合	轻微无力
Ⅳ	中重度	明显无力或不对称	对称，紧张度相同	消失	闭合不全	尽最大力也不对称
Ⅴ	重度	仅感到有运动	不对称	消失	闭合不全	轻微运动
Ⅵ	全瘫	没有运动				

面瘫评定系统（华山版）

项　　目	就诊	针灸前	5 次	10 次	15 次	20 次
1. 眩晕						
2. 眼震						
3. 共济失调						
4. 听力减退						
5. 泪腺分泌减少						
6. 外耳疼痛和疱疹						

(续表)

项　　目	就诊	针灸前	5次	10次	15次	20次
7. 听觉过敏						
8. 乳突部疼痛和压痛						
9. 舌前 2/3 味觉障碍						
10. 唾液减少						
11. 前额皱纹						
12. 含泪征						
13. 紧闭双眼						
14. 鼻翼运动						
15. 安静时鼻唇沟						
16. 鼓腮						
17. 露齿微笑(观察上唇)						
18. 吹口哨(或噘嘴)						
19. 紧闭嘴唇(观察下唇)						
20. 颈阔肌反射						
21. 张大口						

表 14-1-2　面瘫评定系统(华山版)项目意义与评分标准

主干通路	分段和分支	症　状	评　分　标　准
内听道部分	(与位听神经同行)	眩晕	有眩晕=1;无=0
		眼震	有眼震=1;无=0
		共济失调	有共济失调=1;无=0
		听力减退	有耳鸣或听力减退=1;无=0
面神经管部分	膝状神经节	泪腺分泌减少(岩浅大神经)	泪腺分泌明显减少=2;轻度=1;正常=0
		外耳疼痛和疱疹	有外耳疼痛和疱疹=2;外耳疼痛而无疱疹=1;正常=0
	乳突部	听觉过敏(镫骨神经)	明显听觉过敏=2;轻度=1;正常=0
		乳突部疼痛和压痛	明显乳突部疼痛和压痛=2;轻度=1;无=0
	鼓室部	舌前 2/3 味觉障碍(鼓索神经)	味觉完全消失=2;部分减退=1;正常=0
		唾液减少(内脏运动)	唾液明显减少=2;轻度=1;正常=0

（续表）

主干通路	分段和分支	症 状	评 分 标 准
出茎乳孔（颅外部分）	额支支配额肌，眼轮匝肌	前额皱纹	全部消失＝5 100％～75％消失＝4 74％～50％消失＝3 49％～25％消失＝2 25％～1％消失＝1 两侧对称＝0
		含泪征	外溢＝1；无＝0
	颧支支配颧肌，眼轮匝肌	紧闭双眼	眼裂≥3 mm或完全瘫痪＝5 眼裂2～3 mm＝4 眼裂0～2 mm＝3 刚能闭合但眼肌无或轻微抵抗＝2 眼肌有一定抵抗但睫毛征阳性＝1 睫毛征阴性、两侧对称＝0
		鼻翼运动	消失＝1；存在＝0
	颊支支配颊肌，口轮匝肌，口开大肌	安静时鼻唇沟	不对称（消失或变深或变浅）＝1；对称＝0
		鼓腮	有颊部船帆现象＝1；两侧对称＝0
		露齿微笑（观察嘴角）	三齿以上至无运动＝4 三齿范围不对称＝3 二齿范围不对称＝2 一齿范围不对称＝1； 没有不对称＝0
	下颌缘支支配下唇诸肌	吹口哨（或噘嘴吹气）	完全不能做＝5 病侧不能闭合＝4 能闭合但漏气＝3 不漏气但明显不对称＝2 轻微不对称＝1 正常＝0
		紧闭嘴唇（观察下唇）	不对称＝1；没有不对称＝0
	颈支支配颈阔肌	颈阔肌反射	颈阔肌无收缩＝3；轻微收缩＝2；明显收缩＝1；对称（正常）＝0
		张大口	颈阔肌无收缩＝3；轻微收缩＝2；明显收缩＝1；对称（正常）＝0

注：面瘫评定系统（华山版）使用说明：

（1）鼓索神经（舌前2/3味觉障碍）和内脏运动纤维（唾液减少）的支配神经通路比较长，不能准确定位在某一点，但可以定位在整条面神经管。

（2）味觉：用棉签蘸50％NaCl分别测患侧及健侧舌前2/3味觉。若患侧咸味不明显，为味觉减退。

（3）听觉：给患者戴上听诊器，然后敲响512 Hz音叉置于听诊器头部，若患侧耳感觉声音较健侧明显增大或刺痛，为听觉过敏。听觉过敏（过听）的机制是：保持鼓膜紧张的鼓膜张肌受三叉神经分支的翼内神经所支配，镫骨肌受面神经所支配，此两肌呈颉颃关系保持平衡。面神经麻痹

时，镫骨肌发生麻痹，因而，鼓膜张肌相对紧张，鼓膜张力高，微小声音产生强的震动，产生过听现象，见于面神经在镫骨肌分支以上的病变。

(4) 睫毛征：强力闭眼时，正常人睫毛多埋在上下眼睑之中，BELL 面瘫时，则睫毛外露。轻度麻痹时，用力闭双眼，开始时睫毛不对称现象不明显，但短时间后，轻度麻痹侧的睫毛会慢慢显露出来。"眼睑震颤现象"的意义相同，做法是：强力闭双眼，检查者用力扳其闭合的上睑，此时一侧上睑有细微的肌肉痉挛颤动现象，另一侧则没有。

(5) 瞬目运动：BELL 面瘫时不对称，麻痹侧缓慢而且不完全。

(6) 含泪征：因为不能闭眼或闭眼无力，不能眨眼，泪水不能及时入泪囊下排，故泪水长日蓄积在眼睛内，形成含泪征。患者在进食或做咀嚼动作时出现流泪现象。

(7) 鼓腮(颊部船帆现象)：患者在说话时，一侧颊部向外扇动鼓起，如船帆样，即为颊部船帆现象。

(8) 颈阔肌现象：患者张大口，检查者用手触摸其两侧颈部肌肉，观察肌肉收缩情况，如果一侧明显紧张，另一侧无，则为颈阔肌现象阳性。

(9) 颈阔肌反射：嘱患者头用力前屈，检查者一手在患者额部加以阻抗，另一手放在颈阔肌上，健侧颈阔肌收缩，麻痹侧不收缩。

第二节　躯体形式障碍自评量表(SSS)

躯体形式障碍是一种以持久地担心或相信各种躯体症状的优势观念为特征的神经症。患者因这些症状反复就医，各种医学检查阴性和医生的解释均不能打消其疑虑。即使有时存在某种躯体障碍，也不能解释所诉症状的性质、程度或其痛苦与优势观念。经常伴有焦虑或抑郁情绪。尽管症状的发生和持续与不愉快的生活事件、困难或冲突密切有关，但患者常否认心理因素的存在。本障碍男女均有，为慢性波动性病程。

躯体形式障碍的诊断标准如下。

1. 症状标准

(1) 符合神经症的诊断标准。

(2) 以躯体症状为主，至少有下列 1 项：① 对躯体症状过分担心(严重性与实际情况明显不相称)，但不是妄想；② 对身体健康过分关心，如对通常出现的生理现象和异常感觉过分关心，但不是妄想。

(3) 反复就医或要求医学检查，检查结果阴性和医生的合理解释均不能打消其疑虑。

2. 严重标准社会功能受损。

3. 病程标准　符合症状标准至少已 3 个月。

4. 排除标准　排除其他神经症性障碍(如焦虑、惊恐障碍，或强迫症)、抑郁症、精神分裂症、偏执性精神病。临床医生必须仔细检查排除各种躯体疾病才能诊断躯体形式障碍。

躯体形式障碍自评量表(somatoform disorder self-report scale,SSS)(表14-2-1)是笔者于2007年编制的针对躯体形式障碍的自我评估量表。躯体症状自评量表(SSS)的项目来源:"躯体形式症状筛查量表(SOMS)"的所有项目、WHO编制的"神经精神病学临床评估草案(SCAN)"的部分项目、认知和睡眠的项目分解,以及基于文化背景差异的项目添加。SSS共有96项,每个项目分为5个级别,≥3为阳性。"全身乏力或易疲劳、头晕或眩晕、心慌或心悸、健忘、入睡很困难、头痛、后背痛、颈肩痛、口干或咽喉干燥、视力模糊"组成SSS简式,以19/20分为划界分,识别躯体形式障碍的敏感性为88%,特异性为82%。SSS分别由10项、80项和96项组成3种版本,可以分别用于临床初步筛查、严重度判断和国际比较。

表14-2-1　躯体形式障碍自评量表(SSS)

项　　目	严　重　度				
	没有	很轻	中等	偏重	严重
1. 嗅觉减退					
2. 嗅觉改变,如嗅到坏味道					
3. 味觉减退					
4. 味觉改变,如吃到坏味道					
5. 视力模糊,如感到眼前有一片迷雾					
6. 视力丧失(失明)					
7. 视野缺损(某些东西看不见)					
8. 复视(一件东西看成两件东西)					
9. 听力丧失					
10. 耳鸣(耳朵里听到持续而单调的声音)					
11. 口干或咽喉干燥					
12. 嘴巴不能发出说话的声音					
13. 肢体乏力					
14. 全身乏力或易疲劳					
15. 肌肉抽搐或痉挛					
16. 步态不稳或不平衡					
17. 肢体颤抖					
18. 肢体有不由自主的动作					
19. 肢体强直、僵硬、沉重					
20. 头痛					
21. 面孔疼痛					

（续表）

项　　目	严　重　度				
	没有	很轻	中等	偏重	严重
22. 颈肩痛					
23. 后背痛					
24. 前胸痛					
25. 胁肋疼痛					
26. 腹痛					
27. 腰痛					
28. 手脚末端疼痛					
29. 关节疼痛					
30. 肛门痛					
31. 肌肉痛					
32. 全身疼痛					
33. 局部皮肤麻木（比如嘴唇、舌头或肢体发麻）					
34. 皮肤对于疼痛和触摸的感觉丧失					
35. 皮肤肿胀感觉					
36. 怪异的皮肤感觉，如蚁走感、针刺感、紧束感					
37. 皮肤出现红斑					
38. 感觉手足发冷或发热					
39. 皮肤瘙痒（没有皮疹时）					
40. 感觉有一股气从腹部升上来					
41. 脑子里听到持续而单调的声音（如嗡嗡声之类）					
42. 有“筋”吊住感觉					
43. 健忘（记性差）					
44. 语言理解、表达、阅读或书写有困难					
45. 容易分心，不能集中注意力					
46. 不能快速思考或脑子一片空白					
47. 头晕或眩晕					
48. 神志不清					
49. 心慌或心悸					
50. 心前区不适					
51. 间歇性高血压（血压有时高、有时低）					
52. 白天多汗（包括手心易出汗）					
53. 夜间多汗（盗汗）					
54. 头部发热感或沉重感或紧束感					
55. 脸色潮红或发热					

（续表）

项　　目	严　重　度				
	没有	很轻	中等	偏重	严重
56. 胸闷					
57. 轻微用力就感到气急或呼吸困难					
58. 过度换气					
59. 咳嗽（没有感冒和呼吸系统炎症时）					
60. 吞咽困难					
61. 喉咙好像有东西堵住（异物感）					
62. 上腹不适或烧灼感					
63. 胃肠胀气					
64. （眼睛或耳朵出现）幻觉					
65. 食欲下降、体重减轻					
66. 反酸					
67. 嘴巴发苦					
68. 恶心					
69. 呕吐					
70. 食物反流					
71. 打呃					
72. 食物过敏					
73. 肛门经常有液体溢出					
74. 大便困难（便秘）					
75. 反复腹泻					
76. 在没有便秘和腹泻的大便次数增加					
77. 饮食习惯改变					
78. 尿频（小便次数比平时增加）					
79. 尿痛（小便时疼痛）					
80. 尿失禁（排尿不受控制）					
81. 尿潴留（不能排尿）					
82. 性欲减退或性冷淡					
83. （仅限女性回答）月经紊乱					
84. （仅限女性回答）月经过多或过少					
85. （仅限女性回答）怀孕期间持续呕吐					
86. （仅限女性回答）痛经					
87. （仅限女性回答）白带过多					
88. （仅限女性回答）乳房胀满或溢乳					
89. 性交时疼痛					

（续表）

项　　目	严重度				
	没有	很轻	中等	偏重	严重
90. 生殖器或其周围不适感（无外伤或炎症）					
91. （仅限男性回答）勃起或射精困难					
92. 入睡很困难					
93. 睡眠浅，常常夜间中途起床					
94. 在夜晚徘徊、踱步，做不适事宜的活动					
95. 在早晨醒得很早（比你以往的习惯早）					
96. 白天睡得很多					

注：评分标准如下。

1＝没有，在近1周内自觉没有该项症状（问题）。

2＝很轻，在近1周内偶尔有该项症状，对自己的生活和社会功能无实际影响。

3＝中等，在近1周内经常出现该项症状，对自己的生活和社会功能有一定的影响。

4＝偏重，在近1周内频繁出现该项症状，对自己的生活和社会功能有相当程度的影响，希望得到治疗。

5＝严重，在近1周内一直存在该项症状，自觉症状十分严重，对自己的生活和社会功能有严重影响，一定要治疗。

（郭起浩）

参考文献

英文参考文献(以第一作者首字母为序排列)

[1] Banich MT. Cognitive neuroscience and neuropsychology[M]. 2nd ed. Boston New York: Houghton Mifflin Company: 2004.

[2] Baron-Cohen S, Wheelwright S, Hill J, et al. The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism[J], 2001, 42(2): 241-251.

[3] Borroni B, Agosti C, Premi E, et al. The FTLD-modified Clinical Dementia Rating scale is a reliable tool for defining disease severity in Frontotemporal Lobar Degeneration: evidence from a brain SPECT study [J]. European Journal of Neurology, 2010, 17: 703-707.

[4] Brunch MT. Neuropsychology[M]. Boston New York: Houghton Mifflin Company, 1997: 124-398.

[5] Cappa SF, Benke T, Clarke S, et al. EFNS guidelines on cognitive rehabilitation: Report of an EFNS task force[J]. European J Neurol, 2005, 12: 665-680.

[6] Chan AS, Choi MK, Salmon DP. The effects of age, education, and gender on the Mattis dementia rating scale performance of elderly Chinese and American individuals [J]. Journal of Gerontology: Psychological sciences, 2001, 56B(6): 356-363.

[7] Cheung RW, Cheung MC, Chan AS. Confrontation naming in Chinese patients with left, right or bilateral brain damage[J]. J Int Neuropsychol Soc, 2004, 10: 46-53.

[8] Crawford JR, Deary U, Starr J, et al. The NART as an index of prior intellectual functioning: a retrospective validity study covering a 66 year interval[J]. Psychol Med, 2001, 31: 451-458.

[9] Daniel T, Kenneth M, Erik A, et al. Naming dynamic and static actions: neuropsychological evidence[J]. J Physiol Paris, 2008, 102(1-3): 80-94.

[10] Duara R, Loewenstein DA, Greig-Custo MT, et al. Diagnosis and staging of mild

cognitive impairment, using a modification of the clinical dementia rating scale: the mCDR[J]. Int J Geriatr Psychiatry,2010,25(3): 282 - 289.

[11] Dugbartey AT, Townes BD, Mahurin RK. Equivalence of the color trails test and trail making test in nonnative English-speakers[J]. Arch Clin Neuropsychol. 2000 Jul,15(5): 425 - 431.

[12] Faerden A, Nesvag R, Barrett EA, et al. Assessing apathy: The use of the Apathy Evaluation Scale in first episode psychosis [J]. European Psychiatry, 2007, 23: 33 - 39.

[13] Gregory C, Lough S, Stone VE, et al. Theory of mind in frontotemporal dementia and Alzheimer's disease: theoretical and practical implications[J]. Brain, 2002,125: 752 - 764.

[14] Greve KW, Farrell JF, Besson PS. A psychometric analysis of the California card sorting test[J]. Archives of Clinical Neuropsychology, 1995,10(3): 265 - 278.

[15] Heilman KM, Valenstein E. Clinical Neurops; ychology[M]. New York Oxford: Oxford University Preaa, 1993: 141 - 337.

[16] Henry JD, Crawford JR, Philips LH. Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis[J]. Neuropsychologia, 2004, 42: 1212 - 1222.

[17] Hsieh S, Irish M, Daveson N, et al. When one loses empathy: its effect on carers of patients with dementia[J]. Journal of Geriatric Psychiatry and Neurology, 2013, 26(3): 174 - 184.

[18] John D. Greene W, Hodges JR. Identification of famous faces and famous names in early Alzheimer's disease[J]. Brain,1996, 119: 111 - 128.

[19] Jorm AF, Scott R, Jacomb PA. Assessment of cognitive decline in dementia by informant questionnaire[J]. Int J Geriatr Psychiatry, 1989, 4: 35 - 39.

[20] Jorm AF. A short form of the informant questionnaire on cognitive decline in the elderly (IQCODE): development and cross-validation[J]. Psychol Med, 1994. 24(1): 145 - 153.

[21] Katzman, RC, Zhanga YM, Qua OY, et al. A Chinese version of the mini-mental state examination; Impact of illiteracy in a Shanghai dementia survey[J]. Journal of Clinical Epidemiology, 1988,41(10): 971 - 978.

[22] Kelly C, Newton-Howes G. Guide to assessment scales in dementia[M]. London: Current Medicine Group Ltd: 2004.

[23] Lee TM, Cheung CC, Chan JK, et al. Trail making across languages[J]. J Clin Exp Neuropsychol, 2000,22(6): 772 - 778.

[24] Lezak MD, Howieson DB, Loring DW. Neuropsychological assessment[M]. 4th ed.

New York: Oxford University Press: 2004.

[25] Li M, Ng TP, Kua EH, et al. Brief informant screening test for mild cognitive impairment and early Alzheimer's disease[J]. Dement Geriatr Cogn Disord, 2006,21 (5-6): 392-402.

[26] Libon DJ, Bondi MW, Price CC, et al. Verbal serial list learning in mild cognitive impairment: a profile analysis of interference, forgetting, and errors[J]. J Int Neuropsychol Soc, 2011,17(5): 905-914.

[27] Lin KN, Wang PN, Liu CY, et al. Cutoff scores of the cognitive abilities screening instrument, Chinese version in screening of dementia[J]. Dement Geriatr Cogn Disord, 2002,14: 176-182.

[28] Litvan I, Goldman JG, Troster AI, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: movement disorder society task force guidelines [J]. Mov Disord, 2012, 27(3): 349-356.

[29] Marutaa C, Guerreiroa M, de Mendonc A, et al. The use of neuropsychological tests across Europe: the need for a consensus in the use of assessment tools for dementia [J]. European Journal of Neurology, 2011, 18(2): 279-285.

[30] Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment[J]. J Psychiatr Res, 2009,43 (4): 411-431.

[31] Monsch AU, Bondi MW, Salmon DP, et al. Clinical validity of the Mattis dementia rating scale in detecting dementia of the Alzheimer type[J]. Arch Neurol, 1995, 52: 899-904.

[32] Morris JC. The clinical dementia rating (CDR): Current version and scoring rules [J]. Neurology, 1993,43(11): 2412-2414.

[33] O'Bryant SE, Lacritz LH, Hall J, et al. Validation of the new interpretive guidelines for the clinical dementia rating scale sum of boxes score in the national Alzheimer's coordinating center database[J]. Arch Neurol, 2010,67(6): 746-749.

[34] Pearlin LL, Mullan JT, Semple SJ, et al. Caregiving and the stress process: an overview of concepts and their measures[J]. Gerontologist, 1990,30(5): 583-594.

[35] Randolph C, Lansing AE, Ivnik RJ, et al. Determinants of confrontation naming performance[J]. Archives of Clinical Neuropsychology, 1999,14(6): 489-496.

[36] Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease[J]. Am J Psychiatry, 1984, 141: 1356-1364.

[37] Sachdev P, Kalaria R, O'Brien J, et al. Diagnostic criteria for vascular cognitive disorders: a VASCOG statement[J]. Alzheimer Dis Assoc Disord, 2014, 28(3): 206-218.

[38] Saxton J, Snitz BE, Lopez OL, et al. Functional and cognitive criteria produce different rates of mild cognitive impairment and conversion to dementia[J]. J Neurol Neurosurg Psychiatry, 2009, 80: 737-743.

[39] Silva D, Guerreiro M, Maroco J, et al. Comparison of four verbal memory tests for the diagnosis and predictive value of mild cognitive impairment[J]. Dement Geriatr Cogn Disord Extra, 2012, 2: 120-131.

[40] Skinner J, Carvalho JO, Potter GG, et al. The Alzheimer's disease assessment scale-cognitive-plus (ADAS-Cog-Plus): an expansion of the ADAS-Cog to improve responsiveness in MCI[J]. Brain Imaging and Behavior, 2012, 5: 22.

[41] Snowden JS, Thompson JC, Neary D. Knowledge of famous faces and names in semantic Dementia[J]. Brain, 2004, 127: 860-872.

[42] Solomon PR, Hirschoff A, Kelly B, et al. A 7 minute neurocognitive screening battery highly sensitive to Alzheimer's disease[J]. Arch Neurol 1998, 55: 349-355.

[43] Stone VE, Baron-Cohen S, Knight RT. Frontal lobe contributions to theory of mind [J]. Journal of Cognitive Neuroscience, 1998, 10: 640-656.

[44] Strauss E, Sherman E, Spreen O. A compendium of neuropsychological tests[M]. 3rd ed. Oxford university press. 2006: 811-840.

[45] Troyer AK, Moscovitch M, Winocur G, et al. Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease[J]. J Int Neuropsych Society, 1998, 4: 137-143.

[46] Troyer AK, Moscovitch M, Winocur G. The effect of focal frontal-and temporal-lobe lesions on verbal fluency clustering and switching[J]. J Int Neuropsycho Society, 1997, 3: 37-48.

[47] Weintraub S, Salmon D, Mercaldo N, et al. The Alzheimer's disease centers' uniform data set (UDS): the neuropsychological test battery[J]. Alzheimer Dis Assoc Disord, 2009, 23(2): 91-101.

[48] Werheid K, Hoppe C, Thöne A, et al. The adaptive digit ordering test clinical application, reliability, and validity of a verbal working memory test[J]. Archives of Clinical Neuropsychology, 2002, 17: 547-565.

[49] Wilber ST, Carpenter CR, Hustey FM. The six-item screener to detect cognitive impairment in Older Emergency Department Patients[J]. Acad Emerg Med, 2008, 10, 15(7): 613-616.

[50] Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment-beyond controversies, towards a consensus: report of the international working group on mild cognitive impairment[J]. Journal of Internal Medicine, 2004, 256: 240-246.

[51] Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden[J]. Gerontologist, 1980,20(6): 649 - 655.

笔者参与发表的与本书相关的论文(以发表日期由近到远为序排列)

[1] Yang Q, Guo Q, Bi Y. The brain connectivity basis of semantic dementia: a selective review[J]. CNS Neuroscience & Therapeutics, 2015, 21: 784 - 792.

[2] Zhao Q, Guo Q, Liang X, et al. Auditory verbal learning test is superior to Rey-Osterrieth complex figure memory for predicting mild cognitive impairment to Alzheimer's disease[J]. Current Alzheimer Research, 2015,12(6): 520 - 526.

[3] Zhao Q, Robert RO, Ding D, et al. Diabetes is associated with worse executive function in both eastern and western populations: shanghai aging study and Mayo clinic study of aging[J]. Journal of Alzheimer's Disease, 2015,47: 167 - 176.

[4] Ma J, Zhang Y, Guo Q. Comparison of vascular cognitive impairment-no dementia by multiple classification methods[J]. Int J Neurosci, 2014, 8: 1 - 20.

[5] Shu H, Shi Y, Chen G, et al. Opposite neural trajectories of apolipoprotein E ε4 and ε2 alleles with aging associated with different risks of Alzheimer's disease[J]. Cereb Cortex. 2014, 21.

[6] Guo Q, He C, Wen X, et al. Adapting the pyramids and palm trees test and the kissing and dancing test and developing other semantic tests for the Chinese population [J]. Applied Psycholinguistics, 2014,35: 1001 - 1019.

[7] Wang P, Shi L, Zhao Q, et al. Longitudinal changes in clock drawing test (CDT) performance before and after cognitive decline[J]. PLoS ONE, 2014, 9(5): e97873.

[8] Ding D, Zhao Q, Guo Q, et al. Prevalence of mild cognitive impairment in an urban community in China: a cross-sectional analysis of the Shanghai aging study[J]. Alzheimers Dement. 2014, 43(2): 114 - 122.

[9] Lei Y, Li Y, Ni W, et al. Spontaneous brain activity in adult patients with moyamoya disease: a resting-state fMRI study[J]. Brain Res. 2014,10: 27 - 33.

[10] Luo J, Zhu G, Zhao Q, et al. Prevalence and risk factors of poor sleep quality among Chinese elderly in an urban community: results from the Shanghai aging study[J]. PLoS ONE, 2013, 8(11): e81261.

[11] Zhang S, Edwards H, Yates P, et al. Partial mediation role of self-efficacy between positive social interaction and mental health in family caregivers for dementia patients in Shanghai[J]. PLoS ONE. 2013,8(12): e83326.

[12] Guo Q, Zhou Y, Wang CJ, et al. An open-label, nonplacebo-controlled study on cistanche tubulosa glycoside capsules (Memoregain®) for treating moderate

alzheimer's disease[J]. Am J Alzheimers Dis Other Demen, 2013,28(4): 363 - 370.

[13] Zhao Q, Guo Q, Li F, et al. The shape trail test: application of a new variant of the trail making test[J]. PLoS ONE, 2013, 8(2): e57333.

[14] Wang J, Zhang J, Xu L, et al. Cognitive impairments in hashimoto's encephalopathy: a case-control study[J]. PLoS One, 2013,8(2): e55758.

[15] Zhang SY, Edwards H, Yates P, et al. Preliminary reliability and validity testing of a self-efficacy questionnaire for Chinese family caregivers[J]. Aging Ment Health. 2013,17(5): 630 - 637.

[16] Zhao Q, Guo Q, Hong Z. Clustering and switching during a semantic verbal fluency test contribute to differential diagnosis of cognitive impairment[J]. Neurosci Bull. 2013,29(1): 75 - 82.

[17] Zhao Q, Lv Y, Zhou Y, et al. Short-term delayed recall of auditory verbal learning test is equivalent to long-term delayed recall for identifying amnestic mild cognitive impairment[J]. PLoS One. 2012,7(12): e51157.

[18] Mok V, Ding D, Fu J, et al. Transcranial Doppler ultrasound for screening cerebral small vessel disease: a community study[J]. Stroke, 2012,43(10): 2791 - 2803.

[19] Zhao Q, Xiong Y, Ding D, et al. Synergistic effect between apolipoprotein E ε4 and diabetes mellitus for dementia: result from a population-based study in urban China [J]. J Alzheimers Dis, 2012,32(4): 1019 - 1027.

[20] Guo Q, Zhou B, Zhao Q, et al. Memory and Executive Screening (MES): a brief cognitive test for detecting mild cognitive impairment[J]. BMC Neurol, 2012, 11, 12: 119.

[21] Lei X, Guo Q, Zhang J. Mechanistic insights into neurotoxicity induced by anesthetics in the developing brain[J]. Int J Mol Sci. 2012,13(6): 6772 - 6799.

[22] Zhao X, Liu Y, Wang X, et al. Disrupted small-world brain networks in moderate Alzheimer's disease: A resting-state fMRI study [J]. PLoS ONE, 2012, 7 (3): e33540.

[23] Mortimer JA, Ding D, Borenstein AR, et al. Changes in brain volume and cognition in a randomized trial of exercise and social interaction in a community-based sample of non-demented Chinese elders[J]. Journal of Alzheimer's Disease, 2012, 30(4): 757 - 766.

[24] Sun YM, Li HL, Guo QH, et al. The polymorphism of the ATP-binding cassette transporter 1 gene modulates Alzheimer disease risk in Chinese han ethnic population [J]. American Journal of Geriatric Psychiatry, 2012, 20(7): 603 - 611.

[25] Zhang SY, Edwards H, Yates P, et al. Development of self-efficacy questionnaire for

chinese family caregivers[J]. International Journal of Mental Health Nursing. 2012, 21(4): 358 - 365.

[26] Xi Q, Zhao XH, Wang PJ, et al. Functional MRI study of mild Alzheimer's disease using amplitude of low frequency fluctuation analysis[J]. Chinese Medical Journal, 2012, 125(5): 858 - 862.

[27] Wang B, Guo Q, Zhao Q, et al. Memory deficits for non amnestic mild cognitive impairment[J]. Journal of Neuropsychology, 2012,6(2): 232 - 241.

[28] Li HL, Shi SS, Guo QH, et al. PICALM and CR1 variants are not associated with sporadic Alzheimer's disease in Chinese patients[J]. J Alzheimers Dis, 2011,25(1): 111 - 117.

[29] Wang J, Guo Q, Zhou P, et al. Cognitive impairment of mild general paresis of the insane: Alzheimer disease-like pattern[J]. Dement Geriatr Cogn Disord, 2011, 31: 284 - 290.

[30] Wang B, Guo Q, Chen M, et al. The clinical characteristics of 2789 consecutive patients in a memory clinic [J]. Journal of Clinical Neuroscience, 2011, 18: 1473 - 1477.

[31] Lin N, Guo Q, Han Z, et al. Motor knowledge is one dimension for concept organization: Further evidence from a Chinese semantic dementia case[J]. Brain and Language, 2011, 119: 110 - 118.

[32] Mortimer JA, Borenstein AR, Ding D, et al. High normal fasting blood glucose is associated with dementia in Chinese elderly[J]. Alzheimers Dement, 2010,6(6): 440 - 447.

[33] Zhou B, Zhao Q, Teramukai S, et al. Executive function predicts survival in Alzheimer disease: a study in Shanghai[J]. J Alzheimers Dis, 2010,22(2): 673 - 82.

[34] He J, Iosif AM, Lee DY, et al. Brain structure and cerebrovascular risk in cognitively impaired patients: Shanghai community brain health initiative-pilot phase [J]. Arch Neurol, 2010,67(10): 1231 - 1237.

[35] Borenstein AR, Mortimer JA, Ding D, et al. Effects of apolipoprotein E-epsilon4 and-epsilon2 in amnestic mild cognitive impairment and dementia in Shanghai: SCOBHI - P[J]. Am J Alzheimers Dis Other Demen, 2010, 25(3): 233 - 238.

[36] Zhao Q, Zhou B, Ding D, et al. Prevalence, mortality, and predictive factors on survival of dementia in Shanghai, China[J]. Alzheimer Dis Assoc Disord, 2010, 24(2): 151 - 158.

[37] Chen M, Guo Q, Cao X, et al. Preliminary study for the six-item screener to detect cognitive impairment[J]. Neuroscience Bulletin, 2010, 26(4): 317 - 321.

[38] Cao X, Guo Q, Zhao Q, et al. The neuropsychological characteristics and regional cerebral blood flow of vascular cognitive impairment-no dementia[J]. Int J Geriatr Psychiatry, 2010, 25: 1168 - 1176.

[39] Guo Q, Cao X, Zhou Y, et al. Application study of quick cognitive screening test in identifying mild cognitive impairment [J]. Neuroscience Bulletin, 2010, 26 (1): 47 - 55.

[40] Guo Q, Zhao Q, Chen M, et al. A comparison study on mild cognitive impairment with 3 memory tests in Chinese individuals[J]. Alzheimer Dis Assoc Disord, 2009, 23(3): 253 - 259.

[41] Huang Y, Cheng Y, Wu J, et al. Cilostazol as an alternative to aspirin after ischaemic stroke: a randomised, double-blind, pilot study[J]. Lancet Neurol, 2008, 7(6): 494 - 499.

[42] Wang WZ, Jiang B, Wu SP, et al. Change in stroke incidence from a population-based intervention trial in three urban communities in China[J]. Neuroepidemiology, 2007, 28(3): 155 - 161.

[43] Fu JH, Lu CZ, Hong Z, et al. Relationship between cerebral vasomotor reactivity and white matter lesions in elderly subjects without large artery occlusive disease[J]. J Neuroimaging, 2006, 16(2): 120 - 125.

[44] Ding D, Hong Z, Zhao SJ, et al. Long-term disability from acute childhood Japanese encephalitis in Shanghai, China[J]. The American Journal of Tropical Medicine and Hygiene, 2006, 77(3): 528 - 533.

[45] Zhang ZX, Roman GC, Hong Z, et al. Parkinson's disease in China: prevalence in Beijing, Xian, and Shanghai[J]. Lancet, 2005, 365: 595 - 597.

[46] Zhang ZX, Zahner GE, Roman GC, et al. Dementia subtypes in China: prevalence in Beijing, Xian, Shanghai, and Chengdu[J]. Arch Neurol, 2005, 62: 447 - 453.

[47] 师云波，郭起浩，于欢，等. 交替流畅性测验识别轻度认知损害与阿尔茨海默病的效度评价[J]. 中国现代神经疾病杂志，2015，15(7)：560 - 564.

[48] 梁小妞，罗剑锋，赵倩华，等. 不同认知功能量表评估受低教育程度老年人认知功能的分析[J]. 中国临床神经科学，2015，23(4)：423 - 427.

[49] 郭起浩，徐岩. 不同病因所致轻度认知功能损害的神经心理检测比较[J]. 内科理论与实践，2015，10(2)：92 - 94.

[50] 席芊，赵小虎，王培军，等. 遗忘型轻度认知损害患者基于低频振幅的静息态功能磁共振成像研究[J]. 上海交通大学学报(医学版)，2014，34(8)：1162 - 1166.

[51] 梁小妞，罗剑锋，郭起浩，等. 轻度认知功能障碍筛查量表研究进展[J]. 中国临床神经科学，2014，22(6)：704 - 710.

[52] 徐岩,郭起浩(译). 血管性认知障碍的诊断标准：国际血管性行为与认知障碍学会的申明[J]. 神经病学与神经康复学杂志,2014,11(3)：144-154.

[53] 雷宇,郭起浩,李彦江,等. 成人缺血性脑血管病的认知损害特征分析[J]. 中华医学杂志,2014,94(13)：984-989.

[54] 郭起浩. 老年认知障碍常用神经心理测验的应用进展[J]. 生命科学,2014,26(1)：59-63.

[55] 刘艳,丁云龙,钮佳丽,等. 老年人认知功能减退知情者问卷判断阿尔茨海默病患者严重程度的效度[J]. 中国临床心理学杂志,2014,22(6)：1050-1052.

[56] 丁玎,赵倩华,郭起浩,等. 上海城区老年性痴呆患病率调查[J]. 中国临床神经科学,2013,21(1)：19-25.

[57] 李先涛,赵倩华,郭起浩,等. 记忆与执行筛查量表在判断阿尔茨海默病严重度中的作用[J]. 中国临床神经科学,2013,21(5)：548-553.

[58] 张曙映,曾莉,郭起浩. 痴呆家属照顾者自我效能研究进展[J]. 中华老年医学杂志,2013,32(8)：907-909.

[59] 陈科良,赵倩华,郭起浩. 额颞叶退行性变的研究进展与 FTDC 诊断标准介绍[J]. 中华脑血管病杂志(电子版),2012,6(3)：1-5.

[60] 郭起浩,洪震. 血管性认知障碍临床诊断中的若干基本概念问题[J]. 内科理论与实践,2012,7(2)：70-72.

[61] 龚文苹,郭起浩,赵倩华. 老年认知障碍患者的共病状况调查[J]. 中华医学杂志,2012,92(3)：156-159.

[62] 陈科良,赵倩华,郭起浩. 额颞叶退行性的发病机制与治疗[J]. 临床药物治疗杂志,2012,10(5)：51-55.

[63] 杨志秀,郭起浩,赵倩华,等. 轻度认知损害的空间结构能力障碍[J]. 中华神经科杂志,2011,44(11)：759-763.

[64] 林春颖,郭起浩,赵倩华,等. 华山命名测验的编制与效度分析[J]. 中华行为医学与脑科学杂志,2011,20(2)：183-186.

[65] 沈树红,郭起浩,王少石,等. 主诉健忘和轻度认知损害的抑郁状况比较[J]. 中华医学杂志,2010,90(17)：1167-1170.

[66] 陈美蓉,郭起浩,周燕,等. 两种遗忘型轻度认知损害患者的随访比较[J]. 中华神经科杂志,2010,43(5)：351-355.

[67] 郭起浩,洪震. 轻度认知损害的诊断与评估[J]. 中华神经科杂志,2010,43(3)：169-170.

[68] 郭起浩,赵倩华,曹歆轶,等. 中文卡片分类测验的编制和效度检验[J]. 中国现代神经疾病杂志,2010,10(2)：208-212.

[69] 郭起浩,赵倩华,洪震. 轻度认知损害的神经心理学研究进展[J]. 中国现代神经疾病

杂志,2010,10(2):181-186.

[70] 郭起浩,赵倩华,周燕,等.缩短测验时间对轻度认知损害识别力的影响研究[J].中国现代神经疾病杂志,2010,10(3):347-351.

[71] 郭起浩,金丽琳,傅建辉,等.不同类型的血管性认知损害的执行功能障碍研究[J].中华神经科杂志,2009,42(5):314-319.

[72] 周燕,郭起浩,洪震.中文修订版智能筛查在阿尔茨海默病和轻度认知损害评估中的作用[J].中国临床神经科学,2009,17(1):49-53.

[73] 郭起浩,沈树红,赵倩华,等.EC301数字加工和计算量表在早期识别阿尔茨海默病中的作用[J].中华老年心脑血管病杂志,2008,10(6):440-443.

[74] 郭起浩,付建辉,袁晶,等.画钟测验的评分方法研究[J].中华神经科杂志,2008,41(4):234-237.

[75] 袁晶,付建辉,郭起浩,等.遗忘型轻度认知损害和非痴呆血管性认知损害的局部脑血流量比较研究[J].中华神经科杂志,2008,41(10):678-682.

[76] 郭起浩,袁晶,金丽琳.右颞叶萎缩致认知功能缺损一例[J].中华神经科杂志,2008,41(3):210-211.

[77] 陈美蓉,郭起浩,洪震.汉语词语阅读测验的应用研究[J].中国行为医学科学,2008,17(7):656-657.

[78] 沈树红,郭起浩,王少石.视觉和听觉逻辑记忆识别轻度认知障碍的效力比较[J].中国临床心理学杂志,2008,16(2):148-150.

[79] 赵倩华,郭起浩,史伟雄,等.言语流畅性测验在痴呆识别和鉴别诊断中的应用[J].中国临床心理学杂志,2007;(3):241-245.

[80] 吴平,郭起浩,洪震.极简短痴呆筛查测验在识别轻度阿尔茨海默病中的作用[J].神经疾病与精神卫生,2007,7(5):333-336.

[81] 郭起浩,金丽琳,洪震,等.动物流畅性测验在中国老人中的应用[J].中国心理卫生杂志,2007,21(9):622-626.

[82] 郭起浩,虞培敏,赵倩华,等.不同记忆测验识别轻度认知损害的差异研究[J].中华神经科杂志,2007,40(9):610-613.

[83] 孙一忞,郭起浩,袁晶,等.4种流畅性测验的上海社区中老年人的常模分和划界分[J].中国行为医学科学,2007,16(8):714-718.

[84] 郭起浩,孙一忞,袁晶,等.8种执行功能测验的上海社区老年人中的应用[J].中国行为医学科学,2007,16(7):628-631.

[85] 赵倩华,郭起浩,洪震,等.早发和晚发遗忘型轻度认知损害与阿尔茨海默病的神经心理学比较研究[J].内科理论与实践,2007,2(2):98-102.

[86] 郭起浩,金丽琳,吴宇洁,等.躯体症状自评量表的编制与试测研究[J].神经疾病与精神卫生,2007,7(2):91-94.

[87] 郭起浩,孙一忞,袁晶,等. 一例顶叶萎缩患者的神经心理学研究[J]. 中国神经精神疾病杂志,2007,33(2): 59-60.
[88] 郭起浩,孙一忞,虞培敏,等. 听觉词语学习测验的社区老人常模[J]. 中国临床心理学杂志,2007,15(2): 132-135.
[89] 虞培敏,朱国行,郭起浩,等. 拉莫三嗪对癫痫患者认知功能与生活质量的影响[J]. 中华神经科杂志,2007,40(9): 630-633.
[90] 郭起浩,吴平,洪震,等. 一种新的周围性面瘫严重度评估方法及其试测研究[J]. 神经疾病与精神卫生,2006,6(5): 329-332.
[91] 虞培敏,郭起浩,丁玎,等. 癫痫患者认知特点的研究[J]. 中国临床神经科学,2006,14(5): 494-500.
[92] 周燕,陆骏超,郭起浩,等. Rey-Osterriche 复杂图形测验在轻度阿尔茨海默病和轻度认知损害患者中的应用研究[J]. 中国临床神经科学,2006,14(5): 501-504.
[93] 郭起浩,洪震,史伟雄,等. 遗忘型轻度认知损害与极轻度阿尔茨海默病患者临床和心理学特征的鉴别[J]. 中国临床康复,2006,10(26): 4-7.
[94] 陈文利,丁里,郭起浩,等. 重症肌无力的认知功能研究[J]. 中国临床心理学杂志,2006,14(1): 28-30.
[95] 史伟雄,郭起浩,洪震,等. 轻度认知功能损害的听觉词语记忆特点研究[J]. 中华老年医学杂志,2006,25(2): 93-97.
[96] 陆晓晔,郭起浩,范维,等. 改良的简易精神状态检查量表在狼疮认知筛查中的作用[J]. 中华风湿病学杂志,2006,10(6): 355-358.
[97] 郭起浩,史伟雄,洪震,等. 阿尔茨海默病患者的汉字阅读能力研究[J]. 中国临床心理学杂志,2006,14(3): 263-266.
[98] 郭起浩,洪震,史伟雄,等. Boston 命名测验在识别轻度认知损害和阿尔茨海默病中的作用[J]. 中国心理卫生杂志,2006,20(2): 81-85.
[99] 陆骏超,郭起浩,洪震,等. 连线测验(中文修订版)在早期识别阿尔茨海默病中的作用[J]. 中国临床心理学杂志,2006,14(2): 118-121.
[100] 陈霞,张振馨,钱彩韻,等. 卡巴拉汀治疗血管性痴呆的开放性多中心随机对照研究[J]. 中华神经科杂志,2005,38(8): 483-487.
[101] 郭起浩,洪震,吕传真,等. Stroop 色词测验在早期识别阿尔茨海默病中的作用[J]. 中华神经医学杂志,2005,4(7): 701-704.
[102] 郭起浩,陆骏超,洪震,等. 神经心理测验区分退行性与血管性轻度认知功能损害研究[J]. 中国临床心理学杂志,2005,13(4): 405-407.
[103] 于欢,郭起浩,洪震,等. 逻辑记忆在早期识别阿尔茨海默病中的作用[J]. 神经疾病和精神卫生,2005,5(2): 89-91.
[104] 郭起浩,洪震,于欢,等. 轻微认知功能损害快速识别方法的初步研究[J]. 中国临床

心理学杂志，2004，12(1)：60－63.

[105] 郭起浩，洪震，吕传真，等. 记忆障碍病因学分析[J]. 中国现代神经疾病杂志，2004，4(2)：97－100.

[106] 张晓明，徐耕，郭起浩，等. 轻微认知功能损害患者的认知保存与缺损特点分析[J]. 中国心理卫生杂志，2004，18(10)：685－687.

[107] 郭起浩，陈瑞燕，洪震，等. 图片学习测验在识别老年人轻微认知功能损害中的作用[J]. 中国神经精神疾病杂志，2004，30(4)：258－261.

[108] 郭起浩，洪震，吕传真，等. Mattis 痴呆评定量表的效度研究[J]. 中国临床心理学杂志，2004，12(3)：237－238.

[109] 郭起浩，洪震，于欢，等. 语义性痴呆的临床、认知和影像学特征研究——附 3 例报告[J]. 中华神经医学杂志，2004，3(5)：363－348.

[110] 郭起浩，洪震，王蓓，等. 轻微认知功能损害的特点和老年记忆功能问卷的编制[J]. 中国行为医学科学，2003，12(6)：675－678.

[111] 郭起浩，陈瑞燕，洪震，等. 不同文化背景老人认知功能比较[J]. 中国心理卫生杂志，2003，17(11)：734－736.

[112] 郭起浩，洪震，傅建辉，等. 语义性痴呆：一例汉语患者的个案研究[J]. 中国临床心理学杂志，2003，11(4)：253－257.

[113] 郭起浩，王蓓，洪震，等. ADAS 评估阿尔茨海默病的认知功能[J]. 神经疾病与精神卫生，2003，3(4)：251－153.

[114] 郭起浩，吕传真，洪震，等. 皮层下结构卒中后听觉词语记忆损害研究[J]. 中国临床心理学杂志，2002，10(2)：81－84.

[115] 郭起浩，吕传真，洪震，等. 不同部位脑卒中后的筹划执行功能状况研究[J]. 中国行为医学科学，2002，11(2)：129－130.

[116] 郭起浩，吕传真，洪震，等. 内囊-基底节区卒中后的记忆功能及相关因素[J]. 中国心理卫生杂志，2002，16(9)：611－613.

[117] 郭起浩，吕传真，洪震，等. 听觉词语记忆测验在中国老人中的试用分析[J]. 中国心理卫生杂志，2001，15(1)：13－15.

[118] 郭起浩，吕传真，洪震，等. Rey-Osterrieth 复杂图形测验在中国正常老人中的应用[J]. 中国临床心理学杂志，2000，8(4)：205－207.

[119] 郭起浩，金通观，张明园，等. 老年慢性精神分裂症和 Alzheimer 病认知功能及 CT 改变的比较[J]. 中华精神科杂志，1998，31(1)：13－16.

[120] 郭起浩，金通观，张明园，等. 老年慢性精神分裂症和 Alzheimer 病认知功能检查结果比较[J]. 中国心理卫生杂志，1998，12(1)：2－6.

[121] 郭起浩，李柔冰，张明园，等. 正常老人神经心理测验 5 年随访[J]. 中国临床心理学杂志，1998，6(1)：24－28.

[122] 郭起浩,张明园.社区老人 Blessed 痴呆量表在 5 年随访中的应用[J].上海精神医学,1998,10:27-30.

[123] 郭起浩,陈兴时,张明园,等.正常老人认知功能和神经学检查的相关分析[J].中国临床心理学杂志,1997,5(3):139-143.

[124] 郭起浩,李柔冰,张明园,等.神经心理测验与轻度痴呆[J].中国临床心理学杂志,1996,4(3):132-143.

[125] 郭起浩,金通观,张明园,等.Alzheimer 病的神经心理学和有关检查的相关性研究:认知功能和 CT 变化的相关性[J].上海精神医学,1996,8(4):196-199.

[126] 郭起浩,陈兴时,张明园,等.Alzheimer 病的神经心理学和有关检查的相关性研究:认知功能和脑电生理学变化的相关性[J].上海精神医学,1996,8(4):199-202.

[127] 郭起浩,张明圆,Simon D,等.一组评估认知功能的神经心理测验在老人中的应用[J].中国临床心理学杂志,1994,2(3):155-157.

世纪出版

神经心理评估

（第二版）

本书在第一版介绍目前国外常用、临床医师急需了解的30多种神经心理测验的基础上，再次扩充了近一倍数量的测验，力求反映针对老年认知障碍的临床神经心理测验的新进展，内容包括测验的来源、版本演变、操作步骤、信度、效度、正常值、机制研究进展、临床意义和使用注意事项。掌握这些测验有助于临床诊断、治疗效果评估、康复措施制定，对于了解大脑的信息处理机制也有重要意义。

◎责任编辑　萧　烈
◎封面设计　赵　军

www.sstp.cn

上架建议：内科学、神经内科
ISBN 978-7-5478-2936-3
9 787547 829363 >
定价：58.00元
易文网：www.ewen.co